RECEIVED by MSC 4/12/2022 1:07:49 PM

State of Michigan
In the Supreme Court

People of the State of Michigan
Plaintiff-Appellee,

-vs-

Supreme Court No. 162373
Court of Appeals No. 345491
Circuit Court No. 18-000074-01-FC

Dametrius Benjamin Posey
Defendant-Appellant.
_____/

Wayne County Prosecutor
Attorneys for Plaintiff-Appellee
_____

State Appellate Defender Office
Attorneys for Mr. Posey

Document	Page Numbers
Test a Witness's Memory of a Suspect Only Once	1a – 18a
Mugshot Exposure Effect	19a – 40a
The Effect of Repeated Lineups	41a – 59a
Justice Project Eyewitness Identification	60a – 87a
Eyewitness Memory Still Not Common-Sense Table	88a
Comprehensive Evaluation of Show Ups	89a – 115a
Do MI Sentencing Guidelines Meet the Legislature Goals	116a -345a

CREW (Citizens for Racial Equity in Washtenaw) 2020 Final Report	346a – 400a
Resentencing Transcript	401a – 438a

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

APPENDIX A

Test a Witness's Memory of a Suspect Only Once

Check for updates

RECEIVED by MSC 4/12/2022 1:07:49 PM

aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science in the Public Interest
2021, Vol. 22(1S) 1S–18S
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1529100611026259
www.psychologicalscience.org/PSPI
SAGE

Test a Witness's Memory of a Suspect Only Once

John T. Wixted[1], Gary L. Wells[2], Elizabeth F. Loftus[3,4], and Brandon L. Garrett[5]

[1]Department of Psychology, University of California, San Diego; [2]Department of Psychology, Iowa State University; [3]Department of Psychological Science, University of California, Irvine; [4]Department of Criminology, Law and Society, University of California, Irvine; and [5]Duke University School of Law

Abstract

Eyewitness misidentifications are almost always made with high confidence in the courtroom. The courtroom is where eyewitnesses make their *last* identification of defendants suspected of (and charged with) committing a crime. But what did those same eyewitnesses do on the *first* identification test, conducted early in a police investigation? Despite testifying with high confidence in court, many eyewitnesses also testified that they had initially identified the suspect with low confidence or failed to identify the suspect at all. Presenting a lineup leaves the eyewitness with a memory trace of the faces in the lineup, including that of the suspect. As a result, the memory signal generated by the face of that suspect will be stronger on a later test involving the same witness, even if the suspect is innocent. In that sense, testing memory contaminates memory. These considerations underscore the importance of a newly proposed recommendation for conducting eyewitness identifications: *Avoid repeated identification procedures with the same witness and suspect*. This recommendation applies not only to additional tests conducted by police investigators but also to the final test conducted in the courtroom, in front of the judge and jury.

Keywords

eyewitness identification, wrongful convictions, malleability of memory

> No man ever steps in the same river twice, for it's not the same river and he's not the same man.
>
> —Heraclitus

In a court of law, a credible eyewitness who confidently identifies a defendant as the culprit of a crime is often thought to provide direct and powerful evidence of guilt. Indeed, judges have traditionally characterized a courtroom identification as having an "independent" and direct "source" in the witness's memory. Although underappreciated in the legal system, despite being almost universally understood by experimental psychologists, an eyewitness identification in court does *not* provide direct evidence of guilt. Nor is it independently sourced in the witness's memory. Instead, by the time of trial, an eyewitness's memory has almost invariably been contaminated by a variety of factors and is therefore highly error prone. As of today, 375 prisoners have been exonerated by DNA testing, 21 of whom were on death row, and it is now widely understood that eyewitness misidentifications contributed to about 70% of these wrongful convictions (Innocence Project, 2020).

Eyewitness misidentifications typically first occur during early stages of a police investigation (e.g., when a lineup is administered), long before trial. Decades of research have therefore focused on proper methods for conducting lineups so as to minimize initial misidentifications. As the relevant research has accumulated over the years, consensus science-based recommendations about proper eyewitness identification procedures have evolved accordingly. The set of guidelines set forth by

Corresponding Author:
John T. Wixted, Department of Psychology, University of California, San Diego
Email: jwixted@ucsd.edu

Test a Witness's Memory of a Suspect Only Once 2a

RECEIVED by MSC 4/12/2022 1:07:49 PM

Wells et al. (2020) include a new recommendation that is the focus of this article. Specifically, Recommendation 8 is as follows: "Avoid repeated identification procedures with the same witness and suspect." In other words, test a witness's memory for a suspect only once.

Under the right conditions, the first eyewitness identification test can provide reliable information. According to a review of the literature, on an *initial* lineup identification test of *uncontaminated* memory conducted in accordance with current recommendations (i.e., when a pristine procedure is used), confidence can be a reliable indicator of accuracy (Wixted & Wells, 2017). That is, a high-confidence identification implies high accuracy, whereas a low-confidence identification implies low accuracy (the veritable definition of eyewitness reliability). How often pristine conditions prevail in the real world is unknown, but it is known that on the *first* test, eyewitness identification evidence is potentially reliable.

No later test provides more reliable information than the first test because memory is malleable (Davis & Loftus, 2018). That is, like other forms of forensic evidence, memory can be contaminated. Critically, the best chance to test uncontaminated memory is the first test because the very act of testing memory can contaminate it (Steblay & Dysart, 2016). The importance of testing a witness's memory for a suspect only once is hard to overemphasize because the failure to abide by that simple rule might account for a large proportion of the wrongful convictions overturned by DNA evidence (Garrett, 2011). Later, we detail specific research findings and representative real-world cases supporting these claims.

Implementing this newly proposed reform should be simple and straightforward because it involves no special training beyond educating police investigators, prosecutors, and judges about its compelling science-based rationale. The purpose of this article is to do just that. We begin by tracing the growing awareness of the importance of the recommendation to avoid repeated testing by briefly reviewing how consensus science-based guidelines for conducting eyewitness identification procedures have evolved over the years.

The Evolution of Guidelines Pertaining to Eyewitness Identification Procedures

Beginning in 1998, teams of scientists (sometimes working with law enforcement and legal practitioners) have been commissioned to draw up recommendations for conducting eyewitness identification procedures four times. The first guidelines were enumerated in a "white paper" (Wells et al., 1998) commissioned by the American Psychology-Law Society (APLS). That document provided four recommendations for conducting police lineups. A police lineup consists of one suspect and several physically similar fillers. Nowadays, the police typically use photo lineups instead of the live lineups that were once the norm (Police Executive Research Forum, 2013). The four recommendations in that white paper were as follows: (a) the lineup administrator should be blind to the identity of the suspect, (b) the eyewitnesses should be informed that the culprit may or may not be in the lineup, (c) the suspect should not stand out in the lineup (i.e., the lineup should be fair), and (d) a confidence statement should be obtained at the time an identification is made and before any feedback from the police. All of these recommendations remain in force today, but in 1998, the importance of testing memory only once was not yet apparent.

One year later, the National Institute of Justice (NIJ) issued another set of science-based guidelines (Technical Working Group for Eyewitness Evidence, 1999). Whereas the 1998 white paper focused on lineups, per se, the NIJ guidelines were much broader, providing recommendations for creating mug books and composites, for interviewing eyewitnesses, for conducting showups (which involve only the suspect), and for conducting lineups. The lineup recommendations were similar to those in the 1998 white paper, albeit with added specificity on some issues (e.g., the recommendation that at least five fillers be included in a lineup). Still, no mention was made about the special importance of testing a witness's memory for a suspect only once.

In 2013, a committee was appointed by the National Academy of Sciences to provide updated recommendations for eyewitness identification tests (National Research Council, 2014). Some of the new recommendations emphasized system-level issues such as training law-enforcement officers in eyewitness identification procedures and conducting pretrial judicial inquiries into the reliability of the eyewitness evidence. With regard to eyewitness identification procedures per se, they reiterated some of the earlier recommendations and added others, such as the recommendation that the eyewitness identification procedure be videotaped. Critically, they also added a new recommendation that reflected increased awareness of the importance of the initial identification. Specifically, their Recommendation 7 was as follows: "Make Juries Aware of Prior Identifications." In justifying this new recommendation, the committee wrote: "In-court confidence statements may also be less reliable than confidence judgments made at the time of an *initial* [emphasis added] out-of-court identification; as memory fails and/or confidence grows

RECEIVED by MSC 4/12/2022 1:07:49 PM

disproportionately" (p. 110). They also noted that "Eyewitness testimony is a type of evidence where (as with forms of forensic trace evidence) contamination may occur pre-trial" (p. 109). Contamination is the crux of the issue.

The next major development occurred when the APLS commissioned Wells et al. (2020) to update the 1998 white paper in light of what has been learned since that time. There are now nine recommendations, including such new recommendations as conducting an interview before the lineup (in part to warn the witnesses against attempting to identify the culprit on social media and elsewhere) and, as noted above, avoiding repeated identifications with the same witness and same suspect. The overarching reason to avoid repeated tests is that memory is malleable. The essential problem is that on a second test, an individual can look familiar because of the exposure during the first test, even when it is not the right person. Next, we consider how the field came to appreciate that fact and how it leads to the conclusion that law enforcement should avoid testing a witness's memory for a suspect more than one time.

Memory Is Malleable

Concerns about the malleability of memory can be traced back to at least Munsterberg (1908), but a scientific consensus about how easily memories can be modified—or even manufactured outright—did not begin to emerge until the mid-1970s. At that time, Loftus and Palmer (1974) and Loftus et al. (1978) reported the once surprising but now widely accepted finding that something as subtle as the nature of a question posed to an eyewitness can influence what the witness later remembers. Subsequent studies showed that people can even be induced to falsely remember entire events that never happened, such as being lost in a shopping mall as a child (Loftus & Pickrell, 1995) or that they were attacked by a vicious animal (Porter et al., 1999).

The examples summarized above pertain to memory tested by recall (i.e., recollecting details pertaining to a prior event), but eyewitness identification is a *recognition* memory test. As noted earlier, the malleability of memory has proven to be a particularly pernicious force on these tests, considering that many of the DNA exonerations involved eyewitnesses who incorrectly "recognized" the innocent suspect as the culprit. A striking example of memory contamination in the context of recognition memory was reported by Morgan et al. (2013). They conducted a study of military personnel who were confined to a mock prisoner-of-war camp during survival-school training. Each trainee experienced about 30 min of physically confrontational interrogation while alone in a room with an instructor. After the interrogation, the trainee was left alone in an isolation cell. Later, a member of the research team entered the cell and asked questions about the interrogator ("Did your interrogator give you anything to eat?") while showing the participant a photograph of a White male (the "foil"), thereby falsely implying that he was the interrogator.

Next, memory for the interrogator was tested using a nine-person, target-absent, simultaneous photo lineup. The photo lineup contained a picture of the foil but not the actual interrogator (i.e., it was a target-absent lineup in which the foil's face had been differentially familiarized under highly suggestive conditions). Participants who had not been exposed to the foil's face following interrogation identified the foil as the interrogator 15% of the time. By contrast, participants who had been exposed to the foil's face identified the foil as the interrogator a remarkable 84% of the time.

As the example presented above clearly indicates, testing memory with suggestive or otherwise improper procedures contaminates memory. A natural assumption might be that testing memory under optimal conditions (i.e., in accordance with current recommendations) would not have a contaminating effect. This may very well be true when the memory test in question consists of interviewing a witness about their recollection of details about the crime (a recall test) using a proper procedure such as the Cognitive Interview (Fisher & Geiselman, 1992). When questioned properly, witnesses tend to recall accurate information. Recalling accurate information does not contaminate memory. Indeed, it can reinforce it by making the retained information more durable than it otherwise would be, a memory-enhancing phenomenon known as "the testing effect" (e.g., Roediger & Karpicke, 2006).

Unfortunately, the same is not true when memory is tested using a recognition procedure such as a lineup. Even when using a pristine lineup procedure that happens to involve an innocent suspect, testing memory generally contaminates memory for that individual, thereby rendering any later recognition test prejudicial. Next, we consider some theoretical concepts derived from years of basic-science research to understand how and why that happens.

A Primer on the Theoretical Understanding of Recognition Memory

Several long-standing and influential theoretical considerations help to make sense of recognition memory: (a) encoding specificity, (b) similarity-based matching, (c) elaborative processing, (d) signal detection theory, and (e) the source-monitoring framework. These are

RECEIVED by MSC 4/12/2022 1:07:49 PM

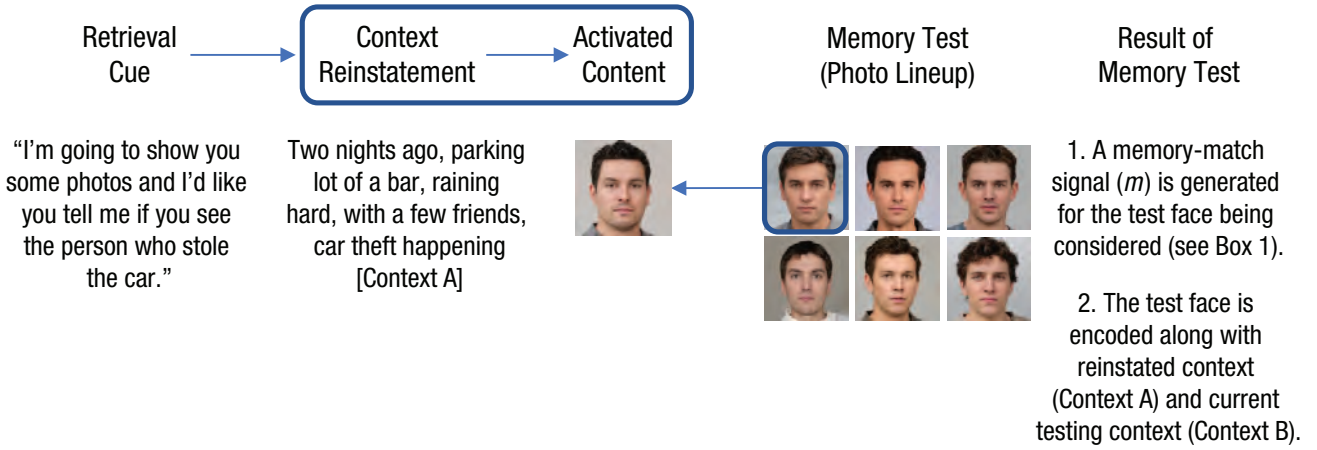


Fig. 1. The specific reference to the witnessed crime by the investigating officer (left column) is a retrieval cue that reinstates the encoding context in the mind of the eyewitness (Context A), which activates the relevant content—the face of the culprit. If multiple culprits were involved, their faces would also be activated, as would faces similar to those of the culprit(s) that might have been seen by the witness in other contexts. We omit those considerations for simplicity. Next, the lineup administrator presents the photo array to the witness, and each photo is compared with the activated content to make an identification decision. This figure illustrates that comparison process taking place for the top-left photo in the lineup. The comparison process yields a memory-match signal associated with the tested face (*m*) that is conceptualized in terms of signal detection theory (see Box 1). After all the faces in the array have been compared with memory, there will be six memory-match signals, and the face associated with the strongest signal will be a candidate for being identified. Comparing a face to memory involves elaborative processing and so incidentally creates a distinctive memory of the tested face (a process illustrated in Fig. 2), one that is encoded along with aspects of the reinstated context (Context A) and the testing context (Context B).

all standard "textbook" ideas that inform our understanding of the intuitively simple but surprisingly complex act of recognizing a once-seen face. The theoretical issues discussed in the remainder of this section are outlined in Figure 1.

Encoding specificity

Memory is generally understood to be cue-dependent (Tulving, 1983; Tulving & Thomson, 1973), which is to say that what you explicitly remember is determined by a *retrieval cue* that distinguishes the sought-after memory from the multitude of memories stored in one's brain. When memory is tested using a lineup, the retrieval cue consists of the specific question put to the witness. This is important because memories are differentially activated and thus accessible depending on the cues available at test (e.g., Godden & Baddeley, 1975). The question posed to the witness is not—or should not be—"are any of these faces familiar?" Instead, the more direct question is: "Do you see the person who committed the crime?" That retrieval cue will reinstate the context of the crime and activate the relevant content (i.e., the face of the culprit) in the brain of the eyewitness, as illustrated in the leftmost columns of Figure 1.

Similarity-based matching

In the simplest and perhaps most common case, the activated content consists of only one face (the singular culprit). If, instead, multiple culprits were involved, all their faces would be activated. According to global matching models, beginning with Gillund and Shiffrin (1984), each recognition-test item (e.g., each face in the lineup) is separately and individually compared against the activated faces (McClelland & Chappell, 1998; Shiffrin & Steyvers, 1997). In the case of a single culprit, this process reduces to what one might already intuitively assume to be true: Each face in the lineup is separately compared against the remembered face of the culprit. Figure 1 illustrates the comparison process for one face in the lineup, which yields a memory-match signal (*m*) for that face. The more similar the face in the lineup is to the witness's memory of the culprit, the stronger this memory signal will be.

Signal detection theory

The memory-match signal (*m*) associated with a tested face is usually conceptualized in terms of signal detection theory. According to this theory, the memory signal is not all or none (i.e., match vs. no match) but is

Test a Witness's Memory of a Suspect Only Once 5a

RECEIVED by MSC 4/12/2022 1:07:49 PM

instead continuous because a face in the lineup can have any degree of similarity to the face of the culprit in memory. It seems natural to assume that, in a target-present lineup, the guilty suspect's face will be the most similar (generating the highest value of *m*), whereas the fillers will be less similar (generating weaker values of *m*). Likewise, in a fair target-absent lineup, no one will be very similar to the face of the culprit in memory, so they should all generate weak values of *m*. These assumptions are sensible, but they omit an important consideration. According to signal detection theory, only *on average* is a guilty suspect expected to generate a stronger memory signal than an innocent suspect or a filler. The reason is that memory matching is an inherently noisy process (Box 1). Thus, occasionally (but not usually), a guilty suspect will generate a weak memory signal and an innocent suspect (or a filler) will generate a strong memory signal. A troubling implication is that, even under ideal conditions involving no memory contamination and pristine testing procedures, and even on the initial test, misidentifications will inevitably happen from time to time. Still, high-confidence misidentifications should be rare. However, for reasons explained next, misidentifications would be expected to increase if memory is tested a second time.

Elaborative processing

The comparison process between a particular face in the lineup and the activated content of the culprit's face in memory does more than simply yield a memory-match signal. It also creates a detailed memory record because of the face processing that occurred during the identification procedure. In a typical lineup, the suspect and the fillers will be physically similar to each other. For example, to be included in the lineup, the face would ideally match the description of the perpetrator provided by the eyewitness (e.g., clean-shaven 20-year-old White male with short dark hair). Because of how lineups are designed, it will not suffice to perform a superficial scan of each face to make an identification decision, such as taking notice only of the shared features. Instead, each face in the lineup must be more thoroughly processed by attending to additional dimensions of the face (Fig. 2).

The act of attending to additional facial dimensions means that the witness has processed some of the unique features that, in configuration, define how a face in the lineup differs from other faces in the population. In other words, by necessity, a face in a lineup is *elaboratively processed* to decide whether this is the person who committed the crime. Such elaborative processing takes place whether the ultimate identification decision is "yes" or "no." Decades of research have established that the more elaboratively a stimulus is processed, the more likely it is to be later remembered (Craik & Tulving, 1975). Why? Craik (2002) put it this way: ". . . a richly elaborate trace will be more differentiated from other episodic records—this greater *distinctiveness* in turn will support more effective recollection in an analogous way to distinctive objects being more discriminable in the visual field" (pp. 306–307).

Elaborative processing creates a memory *incidentally* (i.e., without intention to form a memory). This is, in fact, the essence of the problem associated with testing a witness's memory for a suspect a second time. On that second test, a newly formed memory of the suspect will be accessible, even if the tested suspect is innocent, and the signal generated by the memory of the suspect's face might now be strong.

Source monitoring

The memory of a previously tested face is defined not only by its strength (i.e., by the magnitude of *m*) but also by the memory of the context that accompanied the encoding of the face. Assigning context to the memory signal is known as *source attribution*, and it can be an error-prone process (Johnson et al., 1993). That is, the witness might misattribute the strong memory signal to Source A when, in fact, the face was actually encountered in Source B. Testing memory for the first time using a police lineup almost seems tailor-made for inducing a source misattribution when memory is tested a second time.

An elaboratively processed face encoded during the initial test is not stored in a vacuum. Instead, it is encoded along with aspects of both the internal (i.e., reinstated) context and the external (i.e., testing) context (e.g., Cox & Shiffrin, 2017; Nelson & Shiffrin, 2013). These contexts are labeled Context A and Context B, respectively, in Figure 1. Critically, on the first test, only the culprit's face has been associated with the context of the crime (unlike any filler or any innocent suspect). However, on the second test, more faces will have been associated with that context, including the face of an innocent suspect being tested a second time. When Source A is again reinstated at the time of the second test ("Do you see the person who committed the crime?"), the activated content would now include not just the culprit's face but also the faces that were previously tested, including the face of the innocent suspect.

Typically, when the police conduct a second lineup with the same suspect and same witness, they use a

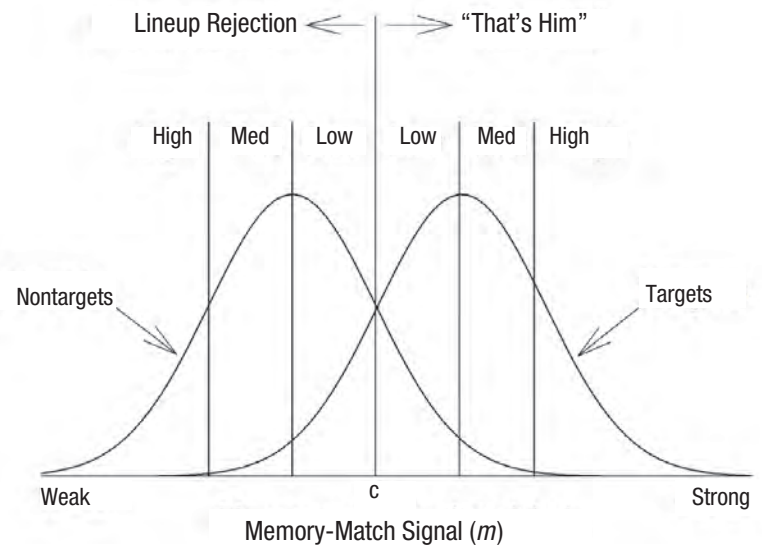
RECEIVED by MSC 4/12/2022 1:07:49 PM

Box 1. Signal Detection Theory

Signal detection theory is a conceptual framework with origins dating back to the dawn of experimental psychology (Fechner, 1860/1966; Green & Swets, 1966; Kellen et al., 2021; Wixted, 2020). As applied to eyewitness identification, signal detection theory conceptualizes the memory-match signal (*m*) that is generated in the brain of an eyewitness when a face in a lineup (innocent or guilty) is compared with the face of the culprit stored in memory. The more similar the two faces are, the stronger the memory-match signal will be. On average, *m* will be strong when the face under consideration is the guilty suspect (the *target*) because that face matches the memory of the culprit, but it will not always be strong. This means that we should think of *m* not as a constant but as a variable that has a range of values across many eyewitnesses who are considering the guilty suspect in a target-present lineup. Its value will be high on average, but it will have variance as well. Thus, in signal detection theory, we represent *m* for guilty suspects as a *distribution* of values with a relatively high mean.

When the face under consideration is an innocent suspect in a target-absent lineup or a filler in either type of lineup (*nontargets*), *m* will be weak, on average, because these faces will not usually be particularly similar to the memory of the culprit. However, it will not always be weak (e.g., the innocent suspect might be a lookalike). Thus, once again, across many lineups and eyewitnesses, it is useful to conceptualize *m* as a variable with a relatively low mean, not as a constant. Thus, there are two distributions (assumed to be Gaussian in form and with equal variance for convenience) with different means, one for guilty suspects and another for both innocent suspects and fillers. For measurement purposes, we can conceptualize the difference between the target and nontarget means in standard deviation units and call that measure *d′* (a value estimated from data in a particular experiment).

After examining all the faces in the lineup, one face will have the maximum value of *m* (m_{max}). If m_{max} is strong enough—that is, if it exceeds the witness's *decision criterion* (c) for making an ID—that face is identified (Wixted et al., 2018). If so, the stronger m_{max} is, the higher the witness's confidence in that ID will be (low, medium or high). If even m_{max} is not strong enough to exceed the witness's decision criterion, the lineup is rejected. The weaker m_{max} is, the higher the witness's confidence will be that the culprit is not present in the lineup. This model inherently predicts a strong confidence–accuracy relationship (Wixted, 2020), which is often observed in lab studies for suspect IDs (Wixted & Wells, 2017) but is less reliably observed for lineup rejections (e.g., Brewer & Wells, 2006).



RECEIVED by MSC 4/12/2022 1:07:49 PM

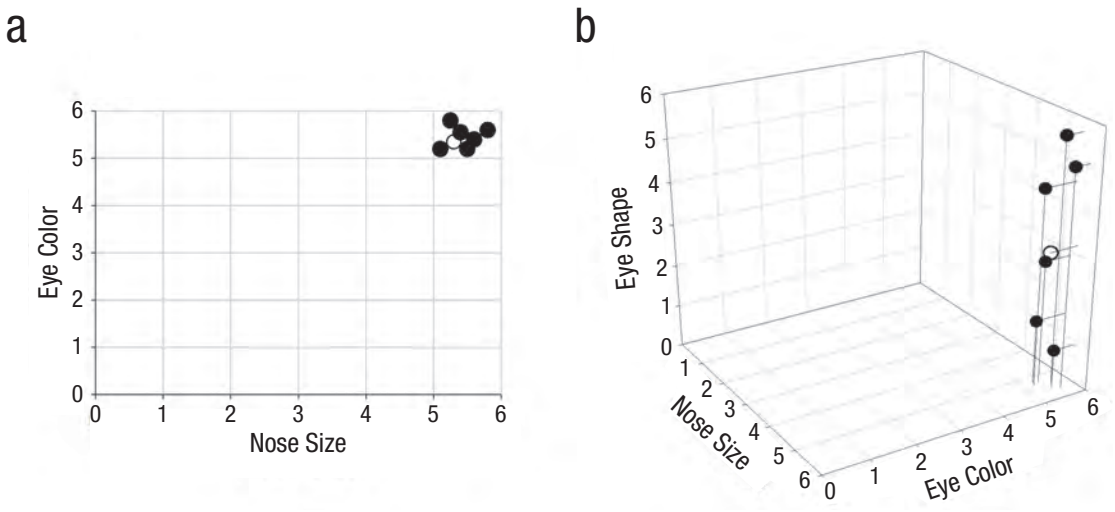


Fig. 2. Multidimensional "face space" (Valentine, 1991; Valentine et al., 2016). Although face perception and memory are widely thought to involve both holistic and feature-based processing (Abudarham et al., 2019; Chua et al., 2015; McKone & Yovel, 2009; Tanaka & Farah, 1993; Tanaka & Simonyi, 2016), low-level perceptual features are used as perceived facial dimensions here for illustrative purposes. The dimensions could just as easily reflect more global properties of any level of abstractness (e.g., masculinity, attractiveness, perceived trustworthiness). The graph in (a) shows perceptual representations of six members of a target-present lineup (filled circles) and the memory representation of the culprit (open circle) along two facial-feature dimensions, nose size and eye color. In this hypothetical example, all values fall between 5 and 6 on both dimensions because the witness described the culprit as having a large nose (0 = *very small* to 6 = *very large*) and dark brown eyes (0 = *very light blue* to 6 = *very dark brown*). The points cluster together because the lineup members were deliberately chosen to match this description of the culprit. When the points cluster together, as they would if only these two features were considered, it is hard for the witness to discriminate the guilty suspect from the fillers. Therefore, additional feature dimensions must be considered. The graph in (b) shows the perceptual and memory representations of the same individuals when a third feature dimension (eye shape) is considered (0 = *round* to 6 = *slanted*). The two-dimensional plot of eye color versus nose size in (a) is now the floor of the three-dimensional plot in (b); the points still cluster together on the floor. The vertical axis represents the new dimension (eye shape). Because this feature was not included in the witness's description, the lineup members exhibit natural variability, so the points spread out along this dimension. Moreover, because eye shape is a feature of the culprit's face that the witness stored in memory but did not describe, only the guilty suspect is now close to the memory representation of the culprit, which generates a differentially strong memory-match signal (*m*). The critical point here is that to make an identification decision, the witness has to consider additional feature dimensions beyond those included in the description. Critically, considering additional feature dimensions individuates a face and is an example of elaborative processing. Elaborative processing makes memories incidentally (i.e., without intention to form a memory).

new set of fillers. Therefore, in the typical case, the suspect will generate a *differentially* strong memory-match signal (potentially attributed to the wrong source) because only that face has been tested previously. The differential familiarization of the suspect's face when memory is tested a second time violates the basic tenets of the "lineups-as-experiments" analogy (Wells & Luus, 1990). The idea is that when police investigators conduct a lineup, they are essentially performing an experiment to test their hypothesis that the suspect is guilty. As experimenters, they should adopt the same protocols that scientists adopt to ensure the integrity of their experiments. One of those protocols is to ensure that participants (witnesses in this analogy) are blind to the hypothesized outcome lest they perform in such a way as to please the experimenter. But if the suspect is the only person in common between the first and the second identification tests, then it is clear to the witness which person the police suspect of having committed the crime—the person in common to both procedures. Because the witness is no longer blind to the suspect's identity on the second test, the lineup is inherently biased against the suspect.

Empirical Studies of Testing Memory a Second Time

In light of the foregoing theoretical considerations, the memory signal generated by the innocent suspect's face will likely be stronger on a second test involving the same witness as a result of the witness having observed the suspect on the first lineup test. The relevant

Test a Witness's Memory of a Suspect Only Once 8a

RECEIVED by MSC 4/12/2022 1:07:49 PM

empirical evidence unambiguously supports this theoretical prediction.

The memory-for-foils paradigm

An illuminating experimental design known as the *memory-for-foils* paradigm provides compelling evidence that testing memory contaminates memory by leaving behind a trace of the tested items (Jacoby et al., 2005). In a typical recognition memory experiment in the basic-science literature, participants are presented with a list of items to study (e.g., a list of words). On a later recognition memory test, those same items (now called "targets") are randomly intermixed with new items ("foils"), and each item is presented individually for a yes/no decision (i.e., "Did this item appear on the list, yes or no?"). Theoretically, the activated memory content against which each test item is compared consists of the items from the study list (Cox & Shiffrin, 2017). In a test like this, the targets are analogous to a guilty suspect because they were seen on the list, and the foils are analogous to innocent suspects and fillers because they were not seen on the list.

After completing the recognition test, the participants are then unexpectedly asked to complete a second recognition memory test consisting of the foils from the first test randomly intermixed with a new set of foils. This time, they are instructed to say "yes" to the foils that appeared on the first test (those items are now the targets) and to say "no" to the new foils. Theoretically, the activated memory set against which test items are compared consists of the items from the just-completed recognition memory test (including the foils). Therefore, the foils, when tested on the surprise memory test, will generate a relatively strong memory-match signal.

Indeed, participants perform very well on that second unexpected test even though, when they first saw the foils (now targets), they were merely attempting to decide whether or not those items had appeared on a previous list, not attempting to memorize them. The foils were elaboratively processed to answer the recognition memory question and were encoded incidentally.

This phenomenon is not limited to lists of words but occurs for faces tested in a lineup as well. In a study reported by Charman and Cahill (2012), participants first viewed a mock-crime video and were later tested using a standard six-person simultaneous photo lineup. Still later, the participants were given a surprise memory test for the five fillers in the lineup. This final test was a list memory test consisting of 10 faces (the five fillers plus five new faces), and each face was presented individually for a yes/no decision about whether it had been seen previously in the lineup. Keep in mind that during the lineup test, the participants were not attempting to memorize the faces; instead, they made only an identification decision about each face. On the final test, the results were striking: The hit rate (percentage of previously seen fillers correctly recognized as such) was 76%, whereas the false alarm rate (percentage of new faces incorrectly recognized as having been previously seen) was only 19%.[1]

It is worth briefly considering how these results are interpreted in terms of the standard signal detection model discussed earlier (Box 1) because it illustrates a key point about how memory contamination caused by testing memory should be conceptualized. The model holds that previously seen faces will generate a stronger memory signal, on average, than new faces. On the unexpected test, the memory signals generated by the new (previously unseen) fillers are conceptualized as having been drawn from a Gaussian distribution with a low mean (Fig. 3a). Because these memory signals are weak, on average, only a small percentage (19%) of them exceed the witness's decision criterion by chance. The memory signal generated by the previously seen fillers (i.e., fillers that became targets) are conceptualized as having been drawn from a Gaussian distribution with a high mean (Fig. 3b). Because these memory signals are strong, on average, a much higher percentage (76%) of them exceed the witness's decision criterion.

For these data, $d' = 1.58$ (Box 1), which means that the participants could easily discriminate previously seen fillers from new fillers. Of most relevance to the issue under consideration here is that this standard theoretical framework conceptualizes the memory distribution of the tested fillers (i.e., fillers that are now targets) as having been shifted upward relative to the fillers that had not yet been tested. In other words, it was not only the 76% of correctly recognized fillers that are now targets that had their memory signals strengthened by the initial test; the remaining 24% were strengthened (i.e., "contaminated") as well but not enough to exceed the decision criterion. Thus, the face was rejected on a second test, but perhaps with less confidence than would otherwise have been the case. The take-home message is that, in theory, testing memory contaminated memory for *all* the tested fillers.

Viewing mug shots

Brown et al. (1977) had participants observe two separate groups of five strangers ("criminals" hereafter). Ninety minutes later, they viewed 15 mug shots, including five people who were criminals and others who were being seen for the first time. A week later, the

RECEIVED by MSC 4/12/2022 1:07:49 PM

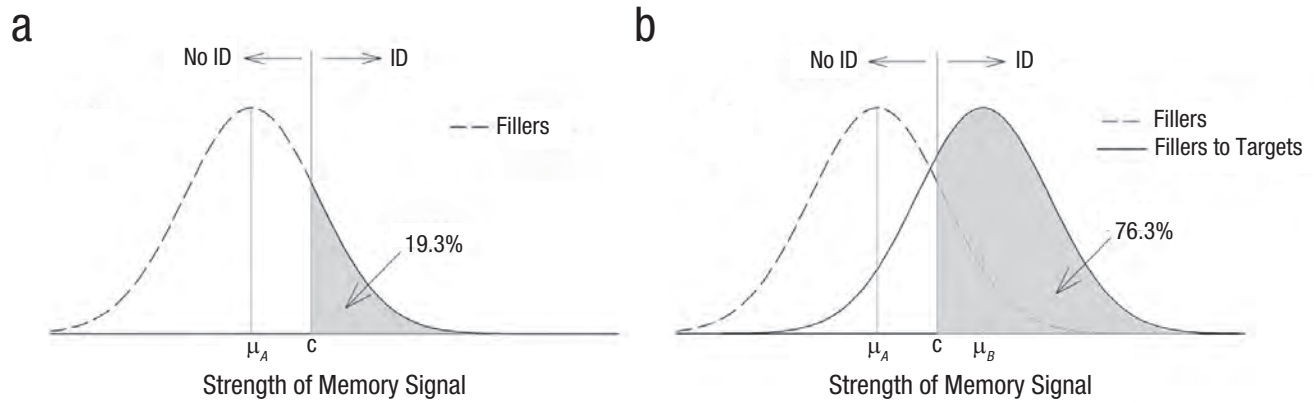


Fig. 3. Signal detection interpretation of findings reported by Charman and Cahill (2012). The graph in (a) represents the standard signal detection interpretation of a false alarm rate of 19.3%. The memory signals generated by new fillers have a mean of μ_A, and the decision criterion (c) is placed well above that. The graph in (b) shows the fillers-to-target distribution with mean μ_B superimposed on the information from (a), showing the standard signal detection interpretation of a hit rate of 76.3%. Note that, relative to new foils, all the fillers that are now targets (i.e., fillers to targets) have had their memory strengths boosted.

participants were presented with lineups (a second memory test) and asked to identify the criminals from the initial in-person encounter. The experience of viewing the mug shots had a clear effect on memory. For lineup members who had never been seen before, the rate of mistaken identification was 8%. However, if a lineup member's mug shot had been seen at the 90-min mark (but not during the original experience), the chances of being falsely identified as a criminal rose to 20%. Thus, the strong memory signals associated with the misidentified mug-shot-only faces were misattributed to the original experience involving the criminals (an example of source misattribution).

The findings reported by Brown et al. (1977) were reinforced by a study reported by Goodsell et al. (2015). Participants watched a short video clip of someone entering an office, after which they were randomly assigned to the mug-shot condition or to the no-mug-shot control condition. Those assigned to the mug-shot condition viewed 50 mug shots of people matched to the description of the culprit from the video, and they were asked to search for the perpetrator. All participants returned after a 48-hr delay and viewed either a target-present lineup or a target-absent lineup.

If a previously seen mug-shot photo in a lineup was one that the participant had previously picked, then that photo was (a) identified as the perpetrator from target-present lineups much more often than a photo of the actual perpetrator (.70 vs. .08) and (b) identified from target-absent lineups with a very high probability (.81). If the previously seen mug-shot photo was *not* the one that the participant had picked, then that photo was (a) *still* identified as the perpetrator from target-present lineups more often than a photo of the actual perpetrator (.28 vs. .18) and (b) identified from target-absent lineups with a probability of .38 (more than double the false-suspect ID rate from the control condition). Thus, memory was contaminated by the initial mug-shot test whether or not the mug-shot face appearing in the later lineup had been previously identified (see related findings reported by Memon et al., 2002).

Testing a suspect a second time

Conceptually similar effects are observed when the initial test consists of viewing a lineup rather than mug shots. In Steblay et al. (2013), participants viewed a video of a simulated crime and then attempted to identify the culprit from two six-person lineups separated by a 2-week retention interval. The suspect (guilty or innocent) was common to both lineups. In the absence of contamination from the first test, the expectation would be that the guilty-suspect ID rate would decline substantially after 2 weeks (because of forgetting) and the false ID rate would remain largely unchanged or increase slightly. For example, with similar retention intervals using a between-subjects design, Palmer et al. (2013) found the correct-ID rate dropped from .60 to .51 (immediate to delayed), $p = .052$, whereas a slight increase in the false-ID rate did not approach significance. By contrast, Steblay et al. (2013) found that when witnesses were tested both immediately and after a delay, the guilty-suspect ID rate *increased* on the delayed test, albeit nonsignificantly (instead of exhibiting the decrease expected as a result of forgetting), and the false-ID rate increased substantially from .21 to .31, $p = .03$. Thus, having seen the suspect in an earlier lineup contaminated memory, placing both innocent

RECEIVED by MSC 4/12/2022 1:07:49 PM

and guilty suspects at greater risk of being identified on a second test than would otherwise be the case.

Testing a witness's memory for a suspect a second time might not be problematic if, on the second test, not only was the same suspect included in the lineup but also the same fillers. In that case, everyone's face would generate an elevated memory signal compared with the first test, and no one would stand out. Lin et al. (2019) conducted this very experiment and found that, even then, nothing was gained by conducting the second test. Instead, witnesses simply became more willing to choose but without improving accuracy.

This Issue Is Specific to Forensic Memory Evidence

In the forensic context, the problem associated with repeated testing is specific to *memory* evidence. For example, repeatedly comparing latent fingerprints lifted from a crime scene to the known fingerprints of a suspect is not problematic and can even serve the cause of justice (e.g., fingerprint examiners can double-check their work) because the test itself does not change the evidence. By contrast, repeated tests of memory are unlikely to serve the cause of justice because testing changes memory (Wells et al., 2020). If comparing latent prints and known prints from a suspect altered the latent prints in such a way as to more closely resemble the fingerprints of the suspect, it seems reasonable to suppose that any fingerprint test after the first would be viewed with suspicion and perhaps excluded from consideration. Although this kind of contamination does not occur with fingerprints, it does occur with "face prints" (the memory of the culprit in the brain of the eyewitness).

Of course, as noted earlier, such contamination can occur even before the first official memory test conducted by the police, so it makes sense to take steps to prevent that from happening. In this regard, Recommendation 1 from Wells et al. (2020) is relevant. The recommendation is to conduct, before the lineup, an interview in which the witness is instructed to avoid attempting to identify the culprit on his or her own. If the witness has already done so and has encountered the suspect's photo (e.g., on social media), thereby contaminating memory before the first official test, it is also important to document that fact.

Memory Contamination Is Not the Only Problem

By focusing on memory contamination resulting from the initial test of memory, we do not mean to imply that it is the only problem associated with testing memory more than once—far from it. For example, as much prior research has shown, the risk to an innocent suspect associated with multiple testing is greatly compounded when suggestive procedures are used and/or when feedback to the witness is provided (Wells & Bradfield, 1998). If, for example, the witness misidentifies the innocent suspect with low confidence from a fair lineup, subsequent feedback from the police can quickly convert it to high confidence (e.g., if the police say "good job, we were pretty sure it was him"). In addition, other memory-contaminating events, such as seeing the face of the suspect again in pretrial hearings or in news stories will further strengthen the memory signal generated by the defendant's face at trial. In addition, if the witness discusses independent evidence against the suspect with prosecutors, it will help to cement the source misattribution according to which the strong memory signal reflects having originally seen the suspect commit the crime.

All of this would be avoided by testing memory only once, thereby strengthening the rationale for the new test-memory-once recommendation in Wells et al. (2020). The new point we are emphasizing here—one that has not received enough attention in the past—is that the witness's memory is *already contaminated* as a result of having taken the first test, even if pristine procedures were followed and even if none of the just-described additional factors exacerbated the problem (as difficult as that might be to imagine).

On the First Test, Confidence Protects Innocent Suspects

On the first (uncontaminated) test using a proper lineup, confidence is more likely to protect than imperil innocent suspects. As noted earlier (Box 1), signal detection theory predicts that decisions made with high confidence should be accurate most of the time, whereas decisions made with low confidence should more often be inaccurate. Related sequential sampling models (e.g., Pleskac & Busemeyer, 2010; Ratcliff, 1978; Ratcliff & Smith, 2004) make similar predictions about reaction time. That is, decisions made quickly should be accurate, whereas decisions made slowly should be less accurate. Empirically, these predictions have often been confirmed in list-memory studies conducted in the basic-science laboratory (e.g., Ratcliff & Murdock, 1976), in lineup studies conducted in the applied-science laboratory (Brewer et al., 2006), and in lineup studies conducted in the real world (Seale-Carlisle et al., 2019). For example, in a study involving actual eyewitnesses to a crime, Seale-Carlisle et al. (2019) reported that lineup decisions made rapidly (e.g., in 5 or 10 s) and with high confidence were estimated to be highly reliable, whereas decisions made slowly (e.g.,

RECEIVED by MSC 4/12/2022 1:07:49 PM

30 s or more) were much less reliable. This was true even in the rare case of a slow decision made with high confidence.

As unfortunate as a misidentification like this would be, keep in mind that the face of the innocent suspect does not actually correspond to the face stored in memory. Therefore, under optimal conditions, the strength of the memory-match signal, despite being randomly strong in a particular case, is not likely to far exceed the witness's decision criterion. According to standard assumptions of signal detection theory (Box 1), and in accordance with much empirical evidence, under ideal testing conditions, misidentifications of the innocent (and of fillers) are usually made with something other than high confidence (Fig. 4).

Losing sight of the low confidence that might be associated with an initial ID (and losing sight of other red flags, such as initial filler identifications or lineup rejections) wastes an opportunity to protect innocent suspects. Identifications made with low confidence are known to be highly error prone, which means that a low-confidence identification should be regarded as an inconclusive test outcome (Wixted & Wells, 2017). This is why the police record of an identification made with low confidence should never be written as "the witness positively identified the suspect." Instead, the record should reflect the lack of confidence, and that lack of confidence should be taken to mean that the memory test was inconclusive. The phrase "positively identified" is probably best reserved for cases where the witness is arguably *positive* that the identified individual is the culprit.

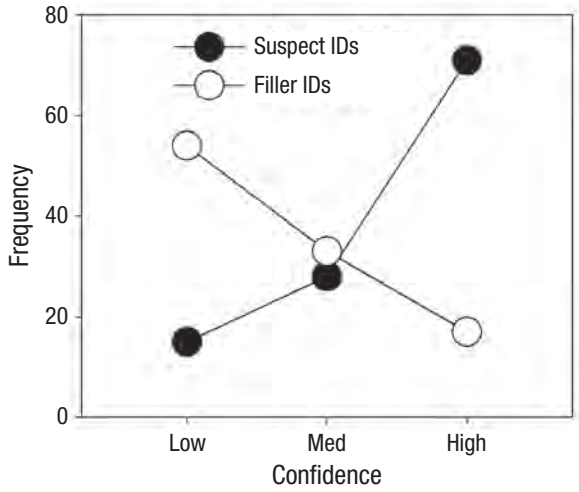


Fig. 4. Number of suspect IDs and filler IDs from 347 photo lineups administered to actual eyewitnesses in the Robbery Division of the Houston Police Department in 2013. The lineups were fair and were administered in double-blind fashion. Of interest here are the IDs made of known innocents (i.e., the fillers), the large majority of which were made with low or medium confidence. Reproduced from Fig. 1b of Wixted, J. T., Mickes, L., Dunn, J. C., Clark, S. E., & Wells, W. (2016). Estimating the reliability of eyewitness identifications from police lineups. *Proceedings of the National Academy of Sciences, USA, 113*, 304–309.

How Should Initial Confidence Be Measured?

The best way to determine whether the witness was "positive" is an actively researched issue, and there is no consensus. Fortunately, the available research suggests that the different methods (e.g., a verbal scale, a 5-point numerical scale, a 100-point numerical scale, asking the witness to use their own words) may not matter very much. For example, Tekin and Roediger (2017) tested 4-, 5-, 20-, and 100-point scales and found that the different scales yielded similar (continuous) confidence-accuracy plots. In their words, "the scales seem convertible from one to the other, and choice of scale range probably does not affect research into the relationship between confidence and accuracy" (p. 2).

Dodson and Dobolyi (2015) considered numerical compared with verbal confidence scales and concluded that "confidence is calibrated with accuracy in a nearly identical manner when confidence is expressed with either a numeric scale or a verbal scale" (p. 267). In agreement with this claim, Tekin et al. (2018) compared 2- and 4-point verbal and numeric scales and found little difference between them. Smalarz et al. (2021) and Mansour (2020) both asked participants to provide confidence in their own words or using a numerical scale. The results indicated that confidence was diagnostic of identification accuracy from a lineup either way, though Mansour (2020) found that verbal statements were more variable.

The upshot of the relevant research is that confidence should be assessed for an initial identification, as has been recommended for many years (Wells et al., 1998). Collecting a confidence statement of some kind appears to be more important than exactly how it is done. Critically, without a confidence statement, it is not possible to know whether the initial ID was made with low confidence, in which case it is highly error prone. Because the initial test is the one that matters, it is essential to collect a confidence statement on that first test. Moreover, the entire identification procedure should be recorded on video (Recommendation 7 of Wells et al., 2020) so that all interested parties—detective, prosecutor, judge, defense counsel, jury, and expert—can see and hear the confidence statement as it was captured in real time.

RECEIVED by MSC 4/12/2022 1:07:49 PM

What if the First Test Involves a Bad Photo of the Suspect?

When a witness fails to identify the suspect or does so with uncertainty on the first test, the police sometimes conclude that the photo of the suspect was not a very good likeness to the face of the suspect. Therefore, they try again, conducting a second test using what they believe to be a better photo. Does a second test endanger an innocent suspect under these conditions?

If the "bad" photo is far from perfect but is nonetheless recognizably the suspect, and if the witness elaboratively processed that bad photo to make a memory-based decision about it (perhaps choosing not to identify that individual), then the witness has processed features that individuate the suspect's face from other faces. The end product of such elaborative processing is an accessible memory record of that face. On the next memory test involving a better photo, those features will match the features that were encoded during the first test. This will have the effect of elevating the strength of the memory signal relative to what it otherwise would have been, thereby imperiling the innocent suspect. In other words, even a bad photo can contaminate memory if it is a recognizable photo of the suspect.

If the bad photo is instead not recognizable as the suspect, so much so that it might as well be a photo of a different person, then it is hard to see how memory contamination would occur. Therefore, in that case, a second test using a better photo would be reasonable. However, whether the photo is recognizable as the suspect is a judgment call. If no suspect ID is made, even a conscientious police investigator who strongly believes that the suspect is guilty might be inclined to honestly conclude that an imperfect-but-recognizable photo of the suspect was "bad," thereby justifying a second test. What can be done to protect against this alluring escape route from our main recommendation to test a witness's memory for a suspect only once?

The best solution would be to preserve the "bad" photo so that others can later judge for themselves whether it is a recognizable photo of the suspect. After all, this is not the only judgment call that an investigating officer has to make. The same officer will have judged the initial photo lineup to be fair, knowing that the photos would be preserved and later judged by others (e.g., by a jury at trial). Preserving the lineup photos incentivizes the investigating officer to exercise caution, ensuring that the lineup is fair. The same principle could be applied to the officer's judgment call about a photo of the suspect being so bad it might as well be a photo of another person. For example, at a pretrial hearing, if the court disagrees with that judgment call, then no later test involving the same suspect and eyewitness should be admissible as evidence.

In addition, the issue can be tested empirically. For example, a sample of people can be presented with the first (allegedly "bad") photo of the suspect and then asked if they can pick that person out from the second lineup. If they cannot do so with greater-than-chance accuracy, then it would be reasonable to conclude that the photo was in fact bad enough that it did not taint the second identification procedure. However, if people can pick the suspect out of the second lineup with greater-than-chance accuracy, then the second lineup should be suppressed.

How Important Was This Issue in the DNA-Exoneration Cases?

Wixted and Wells (2017) argued that on an initial test of *uncontaminated* memory using a *pristine* lineup procedure, high confidence can imply high accuracy, and low confidence can imply low accuracy. Opinions differ as to the reliability of high-confidence IDs in the real world (e.g., Do those tests typically involve uncontaminated memory? Do they typically involve pristine procedures?), but the consensus view is that low-confidence IDs are highly error prone (i.e., at best, they are only weakly probative of guilt). This is true whether or not memory has already been contaminated by the time of the initial test and whether or not pristine procedures are used. Moreover, filler IDs and lineup rejections on the first test are not neutral outcomes but, if anything, are probative of innocence (Wells & Lindsay, 1980).

With that background in mind, consider an analysis reported by Garrett (2011) in his book *Convicting the Innocent*. As noted earlier, data from the Innocence Project show that eyewitness misidentifications contributed to about 70% of more than 375 wrongful convictions later overturned by DNA evidence. Garrett (2011) analyzed the trial records from 161 of those cases in which an eyewitness misidentified an innocent suspect. In the courtroom, at trial, the eyewitness identifications were almost all made with high confidence, which makes sense (otherwise, the prosecutor likely would not have put the witness on the stand). What did these witnesses do at the time of the initial identification? We do not have contemporaneous records, but what these eyewitness and police witnesses described at trial, according to Garrett's (2011) analysis, yielded some interesting observations.

In 57% of the trial transcripts (92 of 161 cases), the witness who misidentified an innocent suspect with

Test a Witness's Memory of a Suspect Only Once 13a

RECEIVED by MSC 4/12/2022 1:07:49 PM

high confidence at trial recalled having initially done so with low confidence (34 cases); they recalled having identified a filler, another suspect, or no one at all (64 cases); or they reported not having seen the culprit's face (15 cases; some cases had more than one type of issue). We do not know what was said at these initial identifications, apart from what the witnesses later recounted at trial. However, to the extent that their recollections are accurate, these initial identifications were highly problematic, not only because the suspect was not confidently identified but for other reasons as well (many of these lineups also involved highly suggestive procedures). This is a problem because IDs made with low confidence are known to be highly error prone. As Garrett (2011) put it, this can provide "a glaring sign that the identification was not reliable" (p. 64). Low-confidence IDs, as well as nonidentifications, filler identifications, or identifications of other suspects, provide an opportunity to protect an innocent yet ultimately misidentified suspect. Unfortunately, for the DNA-exoneration cases involving an inconclusive outcome (or a contrary outcome) on the initial test, that opportunity was lost because the witness's memory was tested more than once.

How many of the remaining cases—the ones for which no testimony about the initial decision exists (43% of 161 cases)—also involved an initial outcome other than a high-confidence ID of the suspect? There is no way to know, in part because in only four cases was the procedure recorded; at a minimum, the evidence reported by Garrett (2011) is consistent with the idea that a sizable fraction of consequential eyewitness misidentifications began with something other than a conclusive (i.e., high-confidence) identification of the suspect. Indeed, as illustrated earlier in Figure 1, it is also consistent with the findings of a police department field study in which misidentifications of known innocents (fillers) were much more likely to have been made with low or medium confidence than high confidence. We turn now to three cases that illustrate how important this issue is.

Three Illustrative Cases

John Jerome White

On August 11, 1979, a man broke into a house in Manchester, Georgia, and raped a 74-year-old woman asleep on her couch (details about this case are available online: National Registry of Exonerations, 2016; "Understanding eyewitness misidentifications," 2011). On the basis of the description of the culprit provided by the victim, the police created a composite sketch. A Georgia Bureau of Investigation agent happened to be investigating a 19-year-old man named Jerome White on another charge, and he thought the sketch resembled White. A week later, the victim picked White out of a photo array, but she was not completely certain (saying she was "almost positive" he was the attacker).



Fig. 5. Live lineup administered to the victim following an initial photo lineup in which she identified Jerome White. White is in the middle, and the actual rapist (James Parham) is the man on the far right.

Perhaps because of those initial signs of uncertainty, she was later administered a live lineup (Fig. 5). White was the only person to appear in both the photo lineup and live lineup (none of the fillers were repeated), so his face had been differentially familiarized as a result of the initial photo-lineup test. Given all the theoretical and empirical considerations we have reviewed to this point, it would not be surprising to learn that the victim identified White again from this lineup, and she did. What makes this case remarkable, however, is who one of the fillers in the lineup turned out to be.

The victim originally told the police that her attacker was "well built" and had a "round face," a description that does not apply to Jerome White (Fig. 5). However, it does apply to one of the fillers in the lineup, namely, the man at the far right in Figure 5. He was not a suspect, but he happened to be in jail at the time, so he was selected to fill out the lineup. Incredibly, many years later, DNA evidence indicated that he was the one who actually committed the rape. Yet after seeing White's face in the initial photo lineup, choosing him, and (evidently) making a source misattribution, he was now the face that came to mind when the victim was asked if she saw the man who raped her in the lineup.

At his trial in 1980, the victim conclusively identified White as the man who had raped her ("that's him"). And why not? She had now seen his face on multiple

Test a Witness's Memory of a Suspect Only Once 14a

RECEIVED by MSC 4/12/2022 1:07:49 PM

previous tests, and the strong sense of familiarity was, in her mind, sourced to the initial crime (not to the lineup tests). The police and prosecutors presumably also reinforced her choice, not to intentionally create an injustice, but to reassure her. Unfortunately, such reassurance serves only to inflate confidence (Wells & Bradfield, 1998). Any doubts the witness had at the initial lineup vanished by the time of the trial. White spent more than 22 years in prison before finally being exonerated by DNA evidence in 2007. The same DNA evidence that exonerated White led prosecutors to charge James Parham (the man on the far right in Fig. 5) with the rape. He pleaded guilty and was sentenced to 20 years in prison.

Steven Gary Titus

On October 12, 1980, Port of Seattle police received a report that a man had raped a female hitchhiker—details about this case can be found from *The Seattle Times* (Henderson, 1981), and from a TED Talk (Loftus, 2013). Steve Titus was a restaurant manager in Seattle at the time, and he was on his way home from a restaurant with his fiancé when his car was stopped by a police officer because it resembled the car that was driven by the rapist. Titus also fit the description of the rapist provided by the victim. When later presented with a photo lineup, the victim identified Titus as her attacker, stating "This one is the closest one." It might very well be the case that Titus provided the closest match to her memory of the culprit (thereby generating the strongest memory signal of the faces in the lineup), but her wording is indicative of low confidence, not high confidence. Yet when Steve Titus was put on trial for rape, the witness's uncertainty had vanished. When she got on the witness stand, she identified Titus with high confidence. By then, not only did the witness have a memory of Titus based (at a minimum) on the initial lineup test, but she may also have been informed of other reasons why police and prosecutors thought he was guilty (further inflating confidence). Moreover, the courtroom identification test itself is inherently suggestive, inflating confidence still further. Based largely on that confident testimony, Titus was found guilty.

According to an article in *The New York Times* (Goleman, 1995), a few months after Titus was convicted, new evidence suggested that a different suspect was responsible for a series of rapes in the area. When the rape victim saw the photograph of the new suspect, she realized that he was the one who had actually raped her. At that point, she began to cry and said "Oh my God, what have I done to Mr. Titus?" However, the key mistake was made by other actors in the criminal-justice system, not the witness, because they tested memory for the suspect (Titus) more than once, ignoring her initial low-confidence ID. Instead of relying on the first test only, they unwittingly relied on contaminated memory evidence at trial to win what turned out to be a wrongful conviction.

Fortunately, Titus avoided a long stint in prison, but the story does not otherwise have a happy ending. Embittered by his wrongful conviction and the financial ruin it caused (including large legal fees and the loss of his job), he decided to file a lawsuit against the Port of Seattle police. Sadly, just before that case was to be heard, Titus died of heart failure at the age of 35.

Charles Don Flores

On the morning of January 29, 1998, witness Jill Barganier saw two people get out of a car outside the home of her neighbor, Elizabeth Black, who was murdered shortly thereafter. Details about this case can be found in legal documents posted at Flores's website (Flores, n.d.). Bargainer described both as White males with long, shoulder-length hair. Another neighbor independently described seeing two White males get out of the car and enter Black's house that morning. When presented with an initial photo lineup containing the main police suspect, a man named Richard Lynn Childs (a White man with long hair down to his shoulders), Barganier immediately identified him with high confidence. Childs owned a handgun of the same caliber used to murder Black, and he owned a conspicuously painted car that multiple eyewitnesses saw parked near Black's home the morning of the murder. He also eventually signed a confession admitting that he shot Black. This was an initial identification made quickly and with high confidence, and all indications are that it was a reliable ID.

Who was the other man Bargainer saw getting out the car that morning? The police suspected Charles Don Flores because he was a known associate of Childs and had been engaged in a drug deal with him in the hours before the murder. Flores was a heavyset Hispanic man with a crew cut and therefore did not match the description of the accomplice provided by the witness. Nevertheless, the police placed his photo in a lineup with other Hispanic men as fillers and presented it to the witness. Quite understandably, the witness did not identify anyone (i.e., she rejected the lineup). This makes sense because it is hard to see why photos of large Hispanic males with short hair would generate a strong memory-match signal compared against the memory of a White male with long hair stored in the witness's brain. Thus, on the initial test, her failure to identify Flores provides no evidence of guilt and instead provides evidence of innocence.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Nevertheless, at the trial, Jill Bargainer was certain that Flores was the man she saw that morning with Childs. Multiple factors presumably contributed to her high confidence, beginning with the memory-based decision she made about Flores on the first test (from that moment on, she likely had a representation of his face in her memory), perhaps continuing with news stories in which his face was shown, and culminating in the suggestive memory test performed at trial. For all these reasons, the only relevant eyewitness evidence was her rejection of the initial lineup.

In addition to the "direct" courtroom evidence provided by the eyewitness, there is indirect evidence against Flores as well. For example, in the days following the crime, he torched the conspicuous paint job on the car driven by Childs the morning of the murder (presumably to make it harder for the police to find), and he fled to Mexico when he learned that the police were looking for him (i.e., he "acted guilty"). This information, if Bargainer were aware of it, would have also served to bolster her confidence by the time of the trial. Childs did not testify about his accomplice at the time of the trial and has not done so to this day.[2]

Despite some independent evidence of guilt, by all accounts, it was the testimony of an extremely credible and highly confident eyewitness that led to the conviction of Flores. In Texas, murder is a capital crime, and an accomplice to a murder is as guilty as the triggerman (Childs). Therefore, Flores was sentenced to death. He has been on death row for over 21 years, and his appeal to the U.S. Supreme Court was denied on January 22, 2021. What may be his final appeal was recently filed in the Texas Court of Criminal Appeals.

The most remarkable fact about this case is that the eyewitness evidence that is mainly responsible for sending him to death row (namely, the witness's confident testimony at trial) is actually probative of innocence when properly understood (i.e., her initial description of the accomplice and her rejection of the initial lineup). In this case, police and prosecutors obviously failed to appreciate that only the first test counts.

A Simple Reform: Test a Witness's Memory for a Suspect Only Once

Presenting the face of a suspect on an eyewitness identification test contaminates the witness's memory for that individual. Such contamination is difficult to avoid, and if it occurs, there is no way to undo it (i.e., there is no way to decontaminate memory). If the witness's memory for that individual suspect is tested again, the suspect's face will generate a stronger memory signal than it otherwise would. As noted earlier, the fact that memory has been contaminated does not necessarily mean that the contaminated memory signal will be strong enough to exceed the criterion for making an identification. However, even in that case, memory has been irretrievably contaminated. Because of source misattribution, witnesses are at risk, on any later test, of responding to the elevated memory signal as if it were based on a memory formed at the time of the crime. By the time of trial, a variety of factors over and above the contaminating effects of testing memory more than once (feedback from the police, seeing the suspects face on the news, etc.) will have likely exacerbated the problem.

In contrast to this science-based perspective, judges often have a different legal perspective. As noted by Garrett (2012), they often embrace the catastrophically mistaken idea that, following the initial test, it is possible to conduct an "independent" test of memory, as if testing the match between a suspect's face and the witness's memory of the culprit multiple times is like testing the match between a suspect's fingerprints and the latent fingerprints lifted from a crime scene multiple times. However, as noted earlier, fingerprints do not change from the first test to the second; memories do. Therefore, once it has been tested and contaminated, it is not possible to perform a second independent test of the memory of a stranger's face that was formed during the commission of the crime.

The only barrier to implementing this recommended reform (testing a witness's memory for a suspect only once) is a faulty theory in the minds of various actors in the criminal-justice system. It therefore follows that implementing this reform should be much simpler than implementing other reforms that require training officers to administer eyewitness identification tests properly. To implement this newly proposed reform, the only training that is required is for policymakers to change their thinking about how memory works and to understand that testing memory for a suspect carries the high risk of irretrievably contaminating memory of that suspect. Considering how many wrongful convictions based on eyewitness misidentification might have been avoided by understanding this simple idea—and considering how many might be avoided going forward—implementing this reform should be an urgent priority.

A Final Word About Courtroom IDs

Because testing memory for a suspect is likely to contaminate memory for that face, a memory test conducted in the courtroom is likely to be a test of contaminated memory, by which time many additional factors exacerbate the problem. There may be rare exceptions (e.g., when the first test of memory for the defendant occurs from the witness stand, at trial, or when the prior test involved a photo that is not recognizable as the

Test a Witness's Memory of a Suspect Only Once 16a

RECEIVED by MSC 4/12/2022 1:07:49 PM

defendant). However, even in cases like that, despite avoiding the problem of memory contamination, a courtroom ID would still be problematic because of its inherently suggestive nature (Wells & Luus, 1990). It is inherently suggestive because only one person is sitting next to the defense attorney, making it plainly obvious to all that prosecutors believe they have enough evidence to be convinced that this is the person who committed the crime.

When it comes to eyewitness identifications, the courts often have it exactly backward, sometimes excluding earlier tests (including the all-important initial test) while allowing in court IDs based on memory that (unbeknownst to the judge) has likely been contaminated by events that occurred after the crime. As Garrett (2012) put it: "Today courts almost always allow courtroom identifications, but they sometimes bar prior identifications. Instead, courts should per se exclude courtroom identifications if there was a prior identification, but they should sometimes admit out-of-court identifications" (p. 457). Perhaps exceptions could be made for rare circumstances such as those mentioned above, but the point is that it makes sense for courts to exclude forensic evidence that has likely been contaminated in a way that is prejudicial to the defendant instead of making an exception for contaminated eyewitness evidence by routinely allowing it.

In addition to excluding courtroom identifications, except under presumably rare circumstances, the only out-of-court identification that should be admitted is the *first* one. Only the first test should be admitted for the same reason the court might exclude other kinds of forensic evidence that is likely to have been contaminated. Barring unusual circumstances (e.g., the witness did even look at the suspect on the first lineup test, or the photo used in the first lineup test was not even recognizable as the suspect), that first test provides the only relevant memory evidence. Even the first official test may involve contaminated memory (e.g., if the witness found a photo of the suspect on social media before viewing the photo lineup), but the first test unarguably provides the best chance to test uncontaminated memory. This simple reform, had it been implemented long ago, could have prevented many (perhaps most) of the wrongful convictions that occurred not because of eyewitness misidentification but because memory was tested more than once.

Transparency

Editor: Nora S. Newcombe

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Notes

1. The false alarm rate was not reported by Charman and Cahill (2012) because the focus of their analysis was different from ours, but the authors kindly provided us with the data.
2. In a plea bargain, Childs was sentenced to 30 years in prison and was released after serving 16 years (i.e., he is a free man today).

References

Abudarham, N., Shkillera, L., & Yovel, G. (2019). Critical features for face recognition. *Cognition*, *182*, 73–83.

Brewer, N., Caon, A., Todd, C., & Weber, N. (2006). Eyewitness identification accuracy and response latency. *Law and Human Behavior*, *30*, 31–50.

Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied*, *12*, 11–30.

Brown, E., Deffenbacher, K., & Sturgill, W. (1977). Memory for faces and the circumstances of encounter. *Journal of Applied Psychology*, *62*, 311–318.

Charman, S. D., & Cahill, B. S. (2012). Witnesses' memories for lineup fillers postdicts their identification accuracy. *Journal of Applied Research in Memory and Cognition*, *1*, 11–17.

Chua, K.-W., Richler, J. J., & Gauthier, I. (2015). Holistic processing from learned attention to parts. *Journal of Experimental Psychology: General*, *144*, 723–729.

Cox, G. E., & Shiffrin, R. M. (2017). A dynamic approach to recognition memory. *Psychological Review*, *124*, 795–860.

Craik, F. I. M. (2002). Levels of processing: Past, present . . . and future? *Memory*, *10*(5-6), 305–318.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*, 268–294.

Davis, D., & Loftus, E. F. (2018). Eyewitness science in the 21st century: What do we know and where do we go from here? In J. Wixted (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience: Vol. 1. Language and memory* (4th ed.). John Wiley & Sons. https://doi.org/10.1002/9781119170174.epcn116

Dodson, C. S., & Dobolyi, D. G. (2015). Misinterpreting eyewitness expressions of confidence: The featural justification effect. *Law and Human Behavior*, *39*, 266–280.

Fechner, G. (1966). *Elements of psychophysics* (Vol. 1). Holt, Rinehart and Winston. (Original work published 1860)

Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The cognitive interview*. Charles C Thomas.

Flores, C. D. (n.d.) *Legal documents*. https://charlesdflores.com/ressources/documents/

Garrett, B. L. (2011). *Convicting the innocent: Where criminal prosecutions go wrong*. Harvard University Press.

Garrett, B. L. (2012). Eyewitnesses and exclusion. *Vanderbilt Law Review*, *65*, 451–506.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*, 1–67.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, *66*, 325–331.

Goleman, D. (1995, January 17). Studies point to flaws in lineups of suspects. *The New York Times*, C1. https://www.nytimes.com/1995/01/17/science/studies-point-to-flaws-in-lineups-of-suspects.html

Goodsell, C. A., Gronlund, S. D., & Neuschatz, J. S. (2015). Investigating mug shot commitment. *Psychology, Crime & Law*, *21*, 219–233.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley.

Understanding eyewitness misidentifications. (2011, March 14). Harvard University Press Blog. https://harvardpress.typepad.com/hup_publicity/2011/03/understanding-eyewitness-misidentifications.html

Henderson, P. (1981, May 15). One man's battle to clear his name. *The Seattle Times*. https://special.seattletimes.com/o/news/local/tituscase/index.html

Innocence Project. (2020). *The causes of wrongful conviction*. https://innocenceproject.org/causes-wrongful-conviction/.

Jacoby, L. L., Shimizu, Y., Daniels, K. A., & Rhodes, M. G. (2005). Modes of cognitive control in recognition and source memory: Depth of retrieval. *Psychonomic Bulletin & Review*, *12*, 852–857.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.

Kellen, D., Winiger, S., Dunn, J. C., & Singmann, H. (2021). Testing the foundations of signal detection theory in recognition memory. *Psychological Review*. Advance online publication. https://doi.org/10.1037/rev0000288

Lin, W., Strube, M. J., & Roediger, H. L. (2019). The effects of repeated lineups and delay on eyewitness identification. *Cognitive Research: Principles and Implications*, *4*, Article 16.

Loftus, E. (2013, March). How reliable is your memory? TED Conferences. https://www.ted.com/talks/elizabeth_loftus_how_reliable_is_your_memory?language=en

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 19–31.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*, 585–589.

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*, *25*, 720–725.

Mansour, J. (2020). The confidence-accuracy relationship using scale versus other methods of assessing confidence. *Journal of Applied Research in Memory and Cognition*, *9*, 215–231.

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review*, *105*, 724–760.

McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic Bulletin Review*, *16*, 778–797.

Memon, A., Hope, L., Bartlett, J., & Bull, R. (2002). Eyewitness recognition errors: The effects of mugshot viewing and choosing in young and old adults. *Memory & Cognition*, *30*, 1219–1227.

Morgan, C. A., Southwick, S., Steffian, G., Hazlett, G. A., & Loftus, E. F. (2013). Misinformation can influence memory for recently experienced, highly stressful events. *International Journal of Law and Psychiatry*, *36*, 11–17.

Munsterberg, H. (1908). *On the witness stand*. McClure.

National Registry of Exonerations. (2016, November 22). *John Jerome White*. https://www.law.umich.edu/special/exoneration/Pages/casedetail.aspx?caseid=3735

National Research Council. (2014). *Identifying the culprit: Assessing eyewitness identification*. The National Academies Press.

Nelson, A. B., & Shiffrin, R. M. (2013). The co-evolution of knowledge and event memory. *Psychological Review*, *120*, 356–394.

Palmer, M. A., Brewer, N., Weber, N., & Nagesh, A. (2013). The confidence-accuracy relationship for eyewitness identification decisions: Effects of exposure duration, retention interval, and divided attention. *Journal of Experimental Psychology: Applied*, *19*, 55–71.

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review*, *117*, 864–901.

Police Executive Research Forum. (2013). *A national survey of eyewitness identification procedures in law enforcement agencies*. https://www.policeforum.org/assets/docs/Free_Online_Documents/Eyewitness_Identification/a%20national%20survey%20of%20eyewitness%20identification%20procedures%20in%20law%20enforcement%20agencies%202013.pdf

Porter, S., Yuille, J. C., & Lehman, D. R. (1999). The nature of real, implanted, and fabricated memories for emotional childhood events: Implications for the recovered memory debate. *Law and Human Behavior*, *23*(5), 517–537.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review*, *83*, 190–214.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*, 333–367.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long term-retention. *Psychological Science*, *17*, 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Seale-Carlisle, T. M., Colloff, M. F., Flowe, H. D., Wells, W., Wixted, J. T., & Mickes, L. (2019). Confidence and response time as indicators of eyewitness identification accuracy in the lab and in the real world. *Journal of Applied Research in Memory and Cognition*, *8*, 420–428.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM–retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*, 145–166.

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

Smalarz, L., Yang, Y., & Wells, G. L. (2021). Eyewitnesses' free-report verbal confidence statements are diagnostic of accuracy. *Law and Human Behavior*, *45*(2), 138-151.

Steblay, N. K., & Dysart, J. E. (2016). Repeated eyewitness identification procedures with the same suspect. *Journal of Applied Research in Memory and Cognition*, *5*, 284–289.

Steblay, N. K., Tix, R. W., & Benson, S. L. (2013). Double exposure: The effects of repeated identification lineups on eyewitness accuracy. *Applied Cognitive Psychology*, *27*, 644–654.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, *46A*, 225–245.

Tanaka, J. W., & Simonyi, D. (2016). The "parts and wholes" of face recognition: A review of the literature. *Quarterly Journal of Experimental Psychology*, *69*, 1876–1889.

Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement* (Report No. NCJ 178240). Offices of Justice Programs, U.S. Department of Justice. http://www.ncjrs.gov/pdffiles1/nij/178240.pdf

Tekin, E., Lin, W., & Roediger, H. L. (2018). The relationship between confidence and accuracy with verbal and verbal + numeric confidence scales. *Cognitive Research: Principles and Implications*, *3*, Article 41. https://doi.org/10.1186/s41235-018-0134-3

Tekin, E., & Roediger, H. L. (2017). The range of confidence scales does not affect the relationship between confidence and accuracy in recognition memory. *Cognitive Research: Principles and Implications*, *2*, Article 49.

Tulving, E. (1983). *Elements of episodic memory*. Clarendon Press.

Tulving, E., & Thomson, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology*, *43*, 161–204.

Valentine, T., Lewis, M. B., & Hills, P. J. (2016). Face-space: A unifying concept in face recognition research. *The Quarterly Journal of Experimental Psychology*, *69*, 1996–2019.

Wells, G. L., & Bradfield, A. L. (1998). "Good, you identified the suspect:" Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology*, *83*, 360–376.

Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., & Wixted, J. T. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law and Human Behavior*, *44*, 3–36.

Wells, G. L., & Lindsay, R. C. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin*, *88*(3), 776–784.

Wells, G. L., & Luus, C. E. (1990). Police lineups as experiments: Social methodology as a framework for properly conducted lineups. *Personality and Social Psychology Bulletin*, *16*, 106–117.

Wells, G. L., Small, M., Penrod, S. J., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, *22*, 603–647.

Wixted, J. T. (2020). The forgotten history of signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*, 201–233.

Wixted, J. T., Mickes, L., Dunn, J. C., Clark, S. E., & Wells, W. (2016). Estimating the reliability of eyewitness identifications from police lineups. *Proceedings of the National Academy of Sciences, USA*, *113*, 304–309.

Wixted, J. T., Vul, E., Mickes, L., & Wilson, B. W. (2018). Models of lineup memory. *Cognitive Psychology*, *105*, 81–114.

Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest*, *18*, 10–65. https://doi.org/10.1177/1529100616686966

RECEIVED by MSC 4/12/2022 1:07:49 PM

APPENDIX B

Mugshot Exposure Effect



RECEIVED by MSC 4/12/2022 1:07:49 PM

Mugshot Exposure Effects: Retroactive Interference, Mugshot Commitment, Source Confusion, and Unconscious Transference

Author(s): Kenneth A. Deffenbacher, Brian H. Bornstein and Steven D. Penrod

Source: *Law and Human Behavior*, Jun., 2006, Vol. 30, No. 3 (Jun., 2006), pp. 287-307

Published by: Springer on behalf of American Psychology-Law Society (AP-LS)

Stable URL: https://www.jstor.org/stable/4499476

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms



Springer and are collaborating with JSTOR to digitize, preserve and extend access to *Law and Human Behavior*

This content downloaded from
141.211.4.224 on Fri, 01 Apon Thu, 01 Jan 1976 12:34:56 UTC
All use subject to https://about.jstor.org/terms

Law and Human Behavior, Vol. 30, No. 3, June 2006 (© 2006)
DOI: 10.1007/s10979-006-9008-1

Mugshot Exposure Effects: Retroactive Interference, Mugshot Commitment, Source Confusion, and Unconscious Transference

Kenneth A. Deffenbacher,[1,4] Brian H. Bornstein,[2] and Steven D. Penrod[3]

Published online: 31 May 2006

More than 25 years of research has accumulated concerning the possible biasing effects of mugshot exposure to eyewitnesses. Two separate metaanalyses were conducted on 32 independent tests of the hypothesis that prior mugshot exposure decreases witness accuracy at a subsequent lineup. Mugshot exposure both significantly decreased proportion correct and increased the false alarm rate, the effect being greater on false alarms. A mugshot commitment effect, arising from the identification of someone in a mugshot, was a substantial moderator of both these effects. Simple retroactive interference, where the target person is not included among mugshots and no one in a mugshot is present in the subsequent lineup, did not significantly impair target identification. A third metaanalysis was conducted on 19 independent tests of the hypothesis that failure of memory for facial source or context results in transference errors. The effect size was more than twice as large for "transference" studies involving mugshot exposure in proximate temporal context with the target than for "bystander" studies with no subsequent mugshot exposure.

KEY WORDS: mugshots; photobiased lineups; retroactive interference; source confusion; unconscious transference.

The phenomenon of the photobiased lineup has existed for as long as police have, in the furtherance of their criminal investigations, engaged in the practice of exposing eyewitnesses to mugshots. If a witness is exposed to mugshots subsequent to viewing a perpetrator and prior to an additional test of recognition memory, there is the possibility that exposure to the mugshot photos may bias the witness's decision at that test, usually a photo or live lineup. Defense attorneys have labeled a photobiased lineup as one in which a possibly innocent person has been arrested because his/her photo was initially included in a set of mugshots but not

[1]Department of Psychology, University of Nebraska at Omaha, Omaha, Nebraska.
[2]Department of Psychology, University of Nebraska-Lincoln, Lincoln, Nebraska.
[3]John Jay College of Criminal Justice, New York.
[4]To whom correspondence should be addressed at Department of Psychology, University of Nebraska at Omaha, Omaha, Nebraska; e-mail: kdeffenbacher@mail.unomaha.edu.

0147-7307/06/0600-0287/0 © 2006 American Psychology-Law Society/Division 41 of the American Psychological Association

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

identified by the witness; subsequently, however, he/she was identified by that witness at a lineup. Any bias thus engendered would be due to the previous exposure of the witness to the mugshot photo of the suspect identified at the subsequent lineup. In such cases, the witness may well have suffered a failure of memory for the circumstances of previous encounter of the face identified at the lineup. This sort of failure of memory for facial source or context is all the more problematic when viewing of the perpetrator has occurred under less than optimal viewing conditions (Brown, Deffenbacher, & Sturgill, 1977). Indeed, the United States Supreme Court had opined that this was the case, even before there was empirical support for their view (*Simmons v. United States*, 1968).

The first empirical examination of the possibility of bias imposed by mugshot exposure more generally as well as the more specific sort of bias that might be engendered by a photobiased lineup was that of Brown et al. (1977). In all three of their experiments, these investigators noted that recognition memory for a once-seen face is much greater than is the capacity to recall the circumstances under which that previously unfamiliar face had been encountered. In Experiments 2 and 3, they employed what we will henceforth refer to as *a transference design* to test for specific biasing effects of mugshot presentation between exposure to live targets and a subsequent lineup.

A transference design occurs in any study in which the ultimate test of recognition memory includes a face of a person who shared proximate temporal context with the target person. Two sorts of proximate temporal context have been investigated with the use of a transference design, one where the "proximate person" was seen in mugshots shortly after encounter with the target person and one where the "proximate person" was encountered as a bystander to the target event. For instance, Read, Tollestrup, Hammersley, McFadzen, and Christensen (1990, Exp. 5) have shown that proximate temporal context for a target face (a sound technician viewed by everyone for 4–5 min in an introductory psychology class) and a "bystander" face (a male bystander who had been seen by half the students in one of several roles, distributing exam materials in other classes, for instance) can be as much as 1 week. That is, they have shown that exposure to target and bystander on the same day is equally productive of transference bias as is exposure to a bystander 1 week after exposure to the target. It could well be that exposure to mugshots as much as 1 week after exposure to the target would be productive of transference bias, too.

At any event, there is a transference effect if a previously seen nontarget face is chosen from a fair lineup, ordinarily a target-absent (TA) one, at a significantly greater rate than for previously unfamiliar foils. Consider a fair target-absent lineup of five persons. If the tendency for witnesses to pick the previously seen innocent person were no greater than the tendency to pick one of the other innocent persons, then 20% of those who had made a selection from the target-absent lineup should have chosen the previously seen innocent. If the actual selection rate were significantly greater than 20%, a transference effect would have been indicated (see Loftus, 1976). It should be noted that some investigators have taken the approach of comparing the rates at which transference errors are made by persons who have been subjected to a transference condition and by persons in a control

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

condition who have not been thus subjected. Clearly, a significantly greater rate of committing a transference error by the experimental group would be indicative of a transference effect. Either approach tests for the presence of a memory bias imposed by a transference condition.

At the lineup, eyewitnesses in Experiments 2 and 3 of Brown et al.'s (1977) study did indeed identify persons seen previously only in mugshots at significantly higher rates than they did previously unfamiliar foils. Interestingly, under Experiment 3's viewing conditions of only a brief glimpse of target persons, witnesses were as likely to indict a person on the basis of a single mugshot encounter as they were a target person encountered only once, that is a target person not also presented in mugshots.

Three other studies employing a transference design involving mugshot exposure have been published subsequent to that of Brown et al. (1977). Deffenbacher, Carr, and Leu (1981) tested whether a transference design would produce significant loss of memory for context of facial encounter at retention intervals of 2 min and 2 weeks. At both retention intervals, they found significantly higher false alarm rates to faces seen once previously in mugshots than to previously unfamiliar foils. More recently, Dysart, Lindsay, Hammond, and Dupuis (2001) provided a specific test of the transference effect and did not obtain a significant effect. Finally, Perfect and Harris (2003) tested for transference effects separately for young and older adults in three separate experiments, providing six different tests of whether a transference design produces negative effects on eyewitness memory. Significant transference effects were obtained in five of the six instances.

It should be noted that the transference effect is not a uniquely forensic phenomenon. In addition to the previously cited studies of eyewitness memory, memory researchers have for the past quarter of a century obtained strong support for a variety of transference-related phenomena. These phenomena include "familiarity without awareness" Mandler (1980), "retention without awareness" (Roediger, 1990), the "false fame effect" (Jacoby, Kelley, Brown, & Jasechko, 1989), and failures in reality monitoring (Johnson & Raye, 1981) and source monitoring (Johnson, Hashtroudi, & Lindsay, 1993). The research literatures associated with each of these related implicit memory phenomena all converge upon the same conclusion: Dissociations between recognition and awareness of context are common.

There is yet another set of studies, not involving mugshot exposure, which nonetheless are relevant to an assessment of whether the phenomenon that has been labeled *unconscious transference* (Williams, 1963) can be reliably produced in studies employing transference designs. Unconscious transference is the term applied to the "transfer of one person's identity to that of another person from a different setting, time, or context" (Read et al., 1990, p. 3). Williams (1963) coined the term in discussing an English murder case, which may have resulted in an innocent person being executed. It seems that one of the eyewitnesses who had identified the defendant had seen him briefly before the crime had occurred and may have engaged in an unconscious transference of identity from the actual perpetrator to the defendant.

Studies using transference designs wherein mugshots are not inserted between viewing of the target and a test of recognition memory, but which do include in that

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

test a face of a bystander who shared proximate temporal context with the target, would appear to be the laboratory analog of the case Williams (1963) has described. Loftus (1976) published the first such study and demonstrated a substantial transference effect. A posthoc one-tailed test performed on her data indicates that the obtained transference effect was not due to the inability of witnesses to discriminate between the target's face and other faces present during the original scenario; the target face was selected from a target-present lineup at a significantly greater rate than was the bystander face from a target-absent one. In testing for presence of the transference effect, Read et al. (1990) conducted five experiments and a total of six independent tests, each involving a TA lineup. Read et al.'s (1990) experiments were field experiments involving the witnessing of bystanders and targets by retail store clerks or by students in university classrooms going about their normal situation-appropriate activities. Tests of the difference between transference and control conditions in regard to the proportion of false positive identifications of the bystander resulted in two negative effects (control condition showing a higher proportion of transference errors), two effect sizes of zero, and two positive effects. Finally, Ross, Ceci, Dunning, and Toglia (1994) presented both transference and control condition witnesses with a film of a staged robbery. Transference witnesses saw an innocent male bystander in the film, whereas control witnesses viewed the same film without the innocent bystander. The former group of witnesses were nearly three times as likely to misidentify the bystander as the perpetrator of the robbery than were the control witnesses.

In addition to the passive effect of familiarity gained by mere mugshot exposure to a previously unfamiliar face, Gorenstein and Ellsworth (1980) sought to determine whether there was a further increase in interference with witness memory for the target face engendered by the witness actively choosing the mugshot of an innocent person.[5] Would witness commitment to a prior choice of a particular mugshot as the target person cause the mugshot image to be retained as strongly or even more strongly than that of the target? In brief, then, a mugshot commitment design involves a control group and an experimental group both being exposed to a target person, followed by mugshot exposure for the latter group. Witnesses in the experimental condition may be urged to pick an innocent person from the mugshots, despite the target's face not being included. At the lineup the target face may or may not be included, but the previously chosen mugshot image definitely is. Gorenstein and Ellsworth (1980) found that whereas control witnesses identified the target from a six-person photo array at a significantly above chance rate (.39), mugshot commitment witnesses failed to identify her at an above chance rate (.22) but did identify their mugshot choice's image at an above chance level (.44), a level somewhat greater than the rate at which control witnesses identified the target person.

Three additional studies employing a mugshot commitment design have been published since that of Gorenstein and Ellsworth (1980). Brigham and Cairns (1988) found the mugshot commitment manipulation both to decrease the identification rate of the actual target (hit rate) and to increase the false alarm rate for the mugshot

[5]Experimental psychologists have used the term *interference* to refer to the tendency for one activity to negatively impact memory for another activity. *Retroactive interference* would reflect a negative impact on memory for material encoded prior to the activity in question.

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

previously identified as the target. Dysart et al. (2001, Exp. 2), using only TA lineups likewise noted a substantial increase in the false alarm rate for witnesses who had made a prior commitment that a particular mugshot "is or maybe" the target. Most recently, Memon, Hope, Bartlett, and Bull (2002) in a slight modification of the mugshot commitment design, had all their witnesses take part in a TA lineup that contained a critical foil (shown prior in the mugshot array) and five new foils. Regardless of their particular choice, choosers of a mugshot from the mugshot array were significantly more likely to choose the critical foil from the TA lineup than were nonchoosers and the control group. This result suggests that the negative effects of mugshot exposure in a mugshot commitment design may not depend entirely on prior commitment to a particular critical foil.

Both the transference and mugshot commitment designs resemble the classic retroactive interference (RI) design used by experimental psychologists, in that both control and experimental conditions are exposed to targets, only the experimental condition is exposed to subsequent mugshots, and witnesses in both conditions are ultimately tested for their memory of the targets. However, both these designs differ from the classic RI design in that they include one or more of the previously seen mugshots in the final test of memory, whereas the RI design does not. The *pure RI design* would include only targets and new, previously unseen foils in a recognition test of memory such as a lineup. Thus, in the case of mugshot studies, the pure RI design would test whether exposure to intervening mugshots suppresses memory for the target items, per se, perhaps due to a form of memory blending in which the identities of a target image and that of a mugshot are confused at encoding (Perfect & Harris, 2003). The transference and mugshot commitment designs, on the other hand, permit assessment of decreases in memory for the target due to a failure of memory for source or context of encounter with a face actually previously seen in mugshots.[6]

Four studies employing the pure RI design with intervening mugshot exposure have been published in the past 25 years. Davies, Shepherd, and Ellis (1979) found a significant suppression of memory for targets when the intervening exposure to mugshots was conducted as a standard mugshot search. On the other hand, Cutler, Penrod, and O'Rourke (1986, Exp. 2), Cutler, Penrod, and Martens (1987), and Dysart et al. (2001, Exp. 1) either obtained decidedly smaller negative effects on witness memory or in the case of Cutler et al. (1987), a positive effect.

Given their widespread use in the furtherance of criminal investigations and their potentially biasing effect having been addressed by the courts (e.g., *Simmons v. United States*, 1968), it has been important for the effects of mugshot exposure to witnesses to have been investigated. More than 25 years of research has accumulated concerning the possible biasing effects of mugshot exposure, research that has generated 32 independent tests of the effects of mugshot exposure on eyewitness memory. In that time, neither the seminal Shapiro and Penrod (1986) metaanalysis of the eyewitness literature nor any subsequent one has been concerned with the effect of mugshot exposure on eyewitness memory. Focused

[6]Deffenbacher et al. (1981) have referred to memory for the target face image(s) per se as item memory and memory for the source or context of encounter with a face as context memory.

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

metaanalyses of other eyewitness phenomena, such as weapon focus (Steblay, 1992), have been conducted after shorter intervals, 15 years in the case of weapon focus, and with as few as 19 relevant effect sizes (Steblay, 1992). Hence, it would be useful to have a firm estimate for both effect size and direction with regard to mugshot exposure, both with regard to proportion correct and the false alarm rate (the rate at which false positive identifications of lineup foils is made). This would allow us to test the hypothesis that mugshot exposure acts both to decrease the proportion of correct identification decisions as well as to increase the false alarm rate.

We also seek to identify variables that might moderate any consistent effect of mugshot exposure on accuracy of memory. In particular, we suspect that type of experimental design (pure RI, transference, or mugshot commitment design) may moderate the effect of mugshot exposure on eyewitness memory.

A final pair of related goals are to determine whether transference designs reliably result in increased rates of transference error and, if so, whether the inclusion of mugshots before the lineup produces any greater rate of transference error than not including mugshots. In other words, we seek to learn whether photobiased lineups are any more productive of transference errors than are lineups resulting from transference situations not involving the potential bias of mugshot exposure, bystander situations, in particular.

METHOD

Sample

Given that the present review was part of a comprehensive metaanalysis project intended to update and extend the one conducted by Shapiro and Penrod (1986), a thorough search of social science citation retrieval systems was conducted. These systems included PsycINfO, Educational Resources Information Center (ERIC), *Sociological Abstracts*, *Dissertation Abstracts International*, Dissertations on-line (http://www.contentville.com/content/dissertations.asp), MEDLINE, and SOCIAL SCISEARCH (the *Social Science Citation Index*). These computer database searches were supplemented by more traditional search methods, including use of bibliographic citations in published research and in social science convention proceedings and contacting leading researchers, in order to identify the most recent published research.

No unpublished studies were included, inasmuch as the legal standards for proffered scientific testimony established by the U.S. Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* (1993) have strengthened the preference by the legal system for metaanalytic conclusions based on a body of well conceived, well executed, and easily retrievable studies. The Court in *Daubert* enunciated four factors that judges should consider in evaluating the admissibility of expert scientific testimony. One factor to consider is whether the theory or technique has been subjected to peer review or publication. Thus, we are in agreement with Clark (2005), Deffenbacher, Bornstein, Penrod, and McGorty (2004), and

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

Reisberg and Heuer (2006) that to survive a *Daubert* review, conclusions about effect size and direction should be based on published studies. We likewise agree with the point made by Clark (2005) that without peer review, unpublished results or the methodological details of the studies which produced them, are not readily available for necessary public scrutiny and evaluation.

Now, certainly the peer review process may have the effect of keeping zero or nearly zero effect sizes, which are more likely to characterize unpublished studies, from accumulating in the scientific literature. However, many unpublished studies have earned their status by having one or more of the following characteristics, low power, ineffective manipulations, severe confounds, or the need for one or more additional experiments to provide adequate controls or to rule out alternative explanations of results. That is, unpublished studies have often either not passed the peer review process or have not been subjected to it, because the investigator has already recognized the problem(s) with the study. There is yet another problem with trying to include unpublished studies. Unlike the situation for published studies, there is no systematic means for gathering unpublished studies. Consequently, one ends up with a set of unpublished papers that is likely not representative of all the extant unpublished papers. The unpublished papers obtained include those one happens to know about, from investigators one happens to know (who have perhaps previously published on the topic—though others might not have), and who happen to have reasonably clean written drafts of their unpublished results and data files. Nevertheless, we shall present a posthoc analysis in the discussion section of this paper, showing that even the weakest of our overall metaanalytic conclusions would not have been changed had we been able to add four additional effect sizes from unpublished studies (Steblay, 1992), in the present case, the exact same effect sizes as those produced in the four unpublished studies in Steblay's sample of 19 studies. These four effect sizes have a mean weighted by sample size of $-.09$.

There were two final study samples. The first sample comprised 11 published articles, providing 32 independent tests of the hypothesis that prior mugshot exposure decreases accuracy on a subsequent lineup. This sample included work published between 1977 and 2003, work yielding a total of 1664 participants.[7] Sample sizes across the 32 tests of the mugshot exposure effect ranged from 26 to 276 ($M = 79.24$).

In order to test whether a reliable unconscious transference effect exists, another sample of seven published articles (1976–2003) was collected, four of which were included in the prior sample, inasmuch as in these prior articles, investigators used a transference design to test for the presence of a mugshot exposure effect. Investigators in the other three articles employed transference designs that did not include mugshot exposure, bystander studies. The entire sample yielded 19 independent tests of the transference effect and a total of 1145 participants, with sample sizes ranging from 24 to 165 ($M = 60.26$).

[7]Even though the cutoff year for the aforementioned comprehensive metaanalysis project was 2002, the Perfect and Harris (2003) article was included for the limited purpose of conducting the present metaanalysis focused just on mugshot exposure effects. This study is one of three published since 2000 and adds six independent tests to a prior total of 26 tests of the hypothesis that mugshot exposure is detrimental to eyewitness memory accuracy. The initial year for the current narrowly focused metaanalyses was determined by the fact that the first published study of mugshot exposure effects occurred in 1977.

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

Finally, it should be noted that even though the current metaanalyses were intended to update and extend the Shapiro and Penrod (1986) metaanalysis, the current sample of studies has little overlap with theirs. The set of studies included in their comprehensive metaanalysis contains only two of 14 articles included in the current sample. These two articles contributed only three of the 32 effect sizes involved in our first two metaanalyses.

Study Characteristics

The previously mentioned comprehensive metaanalysis project, from which the current study springs, has involved the coding of several dozen independent variables for each of 469 existing studies of face recognition from both the eyewitness identification and laboratory face recognition memory traditions. A prior study of interrater reliability at coding 80 of these independent variables showed an average agreement of 93%. Coded variables included stable (e.g., sex and race) and malleable (e.g., disguise) characteristics of both witnesses and targets as well as situational (e.g., target exposure duration) and procedural (e.g., lineup presentation) factors. Variables worthy of mention for their potential usefulness in the present review included type of study (eyewitness identification study or laboratory face recognition study), whether or not a study employed a staged crime, whether the lineup included the target (TP) or not (TA), number of participants, age of participants, number of mugshots inserted between target exposure and subsequent memory test, and obviously, whether mugshot exposure was manipulated. Dependent variables recorded were proportion correct, hit and false alarm rates, if provided, for TP lineups, correct rejection and false alarm rates, when provided for TA lineups, and the signal detection measures, d' and β.

Statistics

To test the statistical reliability of an estimate of the typical effect size found in any particular metaanalysis, we have adopted the Stouffer method (Rosenthal, 1995). Here a metaanalytic Z (Z_{ma}) was calculated by combining Z-scores associated with individual tests of the hypotheses that mugshot exposure and transference designs negatively impact eyewitness memory. The resulting algebraic sum, when divided by $\sqrt{k}$, where k is the number of independent estimates of the effect size, yields the metaanalytic Z. The probability associated with the metaanalytic Z is the overall probability of a Type I error associated with the observed pattern of results. Inasmuch as Z provides an unweighted estimate of the overall probability level, a metaanalytic Z (Z_{mn}) was also calculated which weighted individual Z-scores by the sample size of the study. This allows estimation of population parameters with greater emphasis on larger samples and their more reliable parameter estimates.

It should be noted that whenever recovery of sample sizes and either proportion correct (hits, correct rejections) identifications or false alarm rate per condition permitted, the Z-score entered into the meta-analysis was one calculated for the difference between proportions. When an exact Z-score could not be calculated for a given effect size estimate, a Z-score associated with the p value for the estimate was

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

entered, 1.65 for $p = .05$, for instance. When a test of the hypothesis was reported as not significant, but no statistics were cited, the conservative procedure of entering $Z = 0.00$ was followed (Rosenthal, 1995).

All studywise differences between proportions for experimental (mugshot exposure or transference design) and control conditions (no mugshot exposure or no transference design) were converted into the effect size measure h, inasmuch as h is the coefficient recommended by Cohen (1988) for testing for differences between proportions. Hence, mean effect size for any set of studies and its associated 95% confidence interval is expressed in terms of h.[8]

Finally, given that we have included only published studies in the present metaanalysis, it is quite clear that our sample of studies is not a random sample of all studies that have ever been conducted seeking to test the effects of mugshot exposure. Rosenthal (1995) has pointed out that it is quite likely that published studies have reported lower probabilities of a Type I error than have those studies "squirreled away in file drawers." The concern in regard to this "file drawer problem" is that a sufficient number of studies averaging null results could threaten a metaanalytic conclusion. We have, therefore, adopted Rosenthal's (1995) suggested procedure for calculating a fail-safe N (N_{fs}) in order to determine the number of unknown or not retrieved studies averaging null results required to increase the probability of a Type I error to the just significant level of $p = .05$. Actually, inasmuch as this number is usually a whole number plus a fractional number of studies, we have adopted the rule of rounding to the next higher number. Thus, most values of the fail-safe N that we report in connection with a metaanalytic Z, represent the number of additional null results studies required to increase the probability of a Type I error to a value slightly greater than .05. Clearly, the fail-safe N represents a "tolerance for future null results" (Rosenthal, 1995). We agree with Deffenbacher et al. (2004) that at an absolute minimum the fail-safe N must be at least as large as the number of independent estimates of effect size that went into calculating the metaanalytic Z.

RESULTS

Metaanalysis 1: Mugshot Exposure and Proportion Correct

All Tests

Our first task was to determine the overall status of the hypothesis that mugshot exposure decreases proportion correct (hits, correct rejections) on tests of eyewitness memory for faces. For this analysis, proportion correct for the control condition was subtracted in each instance from that for the experimental condition. Hence, a

[8]Inasmuch as the power to detect a difference between proportions is not simply a function of the magnitude of the difference, but varies as a function of where a given difference occurs along the scale of differences between zero and one, an arcsin transformation (nonlinear) of these differences has been suggested by Cohen (1988). The statistic h is the resulting transform of a given difference in proportions. For instance, the following pairs of proportions all yield approximate values of $h = .20$: .05 and .10, .20 and .29, .40 and .50, .60 and .70, .80 and .87, and .90 and .95.

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

negative metaanalytic Z would indicate a debilitating effect of mugshot exposure on proportion correct performance. The overall proportion of correct identifications at a lineup for the mugshot exposure conditions was .43; the comparable figure for the control conditions was .50 correct. Mean effect size, h, for this analysis was $-.15$ (95% *CI*: $-.33$ to $+.03$); median effect size was $-.17$. The unweighted metaanalytic Z_{ma} was -2.98, $p < .01$, $N_{fs} = 35$. Weighting each of the 15 independent tests of the hypothesis by sample size yielded $Z_{mn} = -2.11$, $p < .025$, $N_{fs} = 10$, a failsafe N that is less than our previously stated standard. Nevertheless, taken together, these analyses provide at least some support for the hypothesis that mugshot exposure inserted between exposure to the target and a subsequent lineup test has a negative effect on proportion correct measures of eyewitness memory.

We next applied a test of the homogeneity of variances across the sample of weighted effect sizes to determine whether the degree of variability exceed that expected on the basis of sampling error alone. A χ^2 value of 84.57, with $df = 14$, $p < .001$, indicated that the degree of variability did indeed exceed that expected on the basis of sampling error. In subsequent analyses, we tested for moderator variables which might account for the considerable variability in the effect sizes for individual tests of the hypothesis (effect size range: $-.694$ to $+.447$, $s = .32$; see Table 1).

Given the relatively small sample of effect sizes and their considerable variability, we were concerned about two potential problems in interpreting results of our moderator analyses. If the variance in effect sizes were "carved up" in several overlapping ways, multicollinearity could be a problem. Further, where there were only a few effect sizes of considerable variability associated with a given moderator variable, it would be possible for one or two large effect sizes to produce a significant metaanalytic Z, but for the 95% *CI* to include zero.

To mitigate these concerns, we took two decisions. First, we decided that we would limit each moderator analysis to two moderator variables, two variables most directly germane to studies of mugshot exposure, the number of mugshots exposed and the nature of the experimental design for producing and testing for effects of mugshot exposure. That is, we limited ourselves to carving the effect size variance in only two ways. Further, if analysis were to reveal a clear confounding of the two moderator variables, we decided that we would then report only the results for one, the one producing the more trustworthy effect size. We defined a *trustworthy effect* for present purposes as one that was associated with a statistically significant metaanalytic Z, a 95% *CI* that did not include zero, and a failsafe N that was at least twice as great as the number of effect sizes in the sample or subsample.

Moderator Analyses

Inasmuch as four of the 15 tests of the hypothesis that mugshot exposure decreases proportion correct were studies wherein the target(s) appeared in mugshots as well as in the final test of memory (Brown et al., 1977, Exp. 2–3; Lindsay, Nosworthy, Martin, & Martynuck, 1994, Exp. 1, both target present and target absent lineups), one might therefore predict not a negative effect of mugshots in these instances, but rather no effect or perhaps a positive effect. That is, having

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

Table 1 Identification Accuracy Effect Sizes (Studies Ordered by Date)

Author	*N*	Proportion correct (*h*)	False alarms (*h*)
Loftus 1976[a]	25		1.26
Brown et al., 1977, Exp. 2	64	.28	.35
Brown et al., 1977, Exp. 3	146	.45	.26
Davies et al., 1979	60	−.45	
Gorenstein & Ellsworth, 1980	36	−.37	.60
Deffenbacher et al., 1981, 2 min	27	−.44	.77
Deffenbacher et al., 1981, 2 week	26	−.24	.41
Cutler et al., 1986, Exp. 2	276	−.17	
Cutler et al., 1987	165	.21	
Brigham & Cairns, 1988	99	−.69	.94
Read et al., 1990, Exp. 1[a]	31		.00
Read et al., 1990, Exp. 2[a]	20		−.61
Read et al., 1990, Exp. 3[a]	69		−.43
Read et al., 1990, Exp. 4[a]	49		.00
Read et al., 1990, Exp. 5A[a]	96		.14
Read et al., 1990, Exp. 5B[a]	131		.34
Lindsay et al., 1994, Exp. 1, TP	69	.00	.00
Lindsay et al., 1994, Exp. 1, TA	70	.00	.00
Ross et al., 1994, Exp. 1[a]	165		.82
Dysart et al., 2001, Exp. 1	104	−.08	
Dysart et al., 2001, Exp. 2[b]	88	−.37	.87
Dysart et al., 2001, Exp. 2[c]	88	.04	−.32
Memon et al., 2002	112	−.46	.46
Perfect & Harris, 2003, Exp. 1A[d]	30		.07
Perfect & Harris, 2003, Exp. 1B[e]	51		.74
Perfect & Harris, 2003, Exp. 2A[d]	45		.64
Perfect & Harris, 2003, Exp. 2B[e]	48		.35
Perfect & Harris, 2003, Exp. 3A[d]	30		.58
Perfect & Harris, 2003, Exp. 3B[e]	30		1.44

[a]Transference studies without mugshot exposure.
[b]Mugshot commitment design.
[c]Transference design with mugshot exposure.
[d]Young adult participants.
[e]Older adult participants.

an extra exposure of the target under ideal viewing conditions might well increase the accuracy of eyewitness memory on a subsequent test, rather than decrease it. Such an effect would impose another sort of bias on any subsequent lineup, especially if the eyewitness had not identified the actual perpetrator from the mugshot presentation. If the perpetrator's image being present in the mugshots increased the probability that he would subsequently be identified at the lineup, justice would have been served, but one of the assumptions of the law concerning eyewitness testimony would not have been met; the perpetrator would have been identified not based strictly on the eyewitness's uncontaminated memory of the crime scene.

The proportion correct for the mugshot exposure conditions, taken across all four aforementioned studies, was .55 versus a figure of .48 for the control conditions. The metaanalytic Z was statistically reliable, $Z_{ma} = 2.33$, $p < .01$, $N_{fs} = 4$. Inasmuch as the confidence interval associated with the average effect size of $h = +.18$ included zero (95% *CI*: −.17 to +.53), and the failsafe N did not meet our criterion for a trustworthy effect, it would be safest to conclude that there is no biasing effect

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

on proportion correct induced by including the target(s) in mugshots as well as a subsequent lineup.

Desiring a measure of possibly biasing effects of mugshot exposure, without the contamination of target presence in the mugshots, a moderator analysis was conducted on the 11 tests of the mugshot exposure hypothesis in which targets did not subsequently appear in mugshots. Overall proportion correct for the mugshot exposure conditions in this instance was .48, with the comparable figure for the control conditions being .61, $Z_{ma} = 4.91$, $p < .0001$, $N_{fs} = 87$. This uncontaminated test of possible mugshot exposure bias yielded an effect size nearly twice as strong as that produced by the preceding omnibus test, $h = -.28$ (95% *CI*: $-.11$ to $-.45$).

Further moderator analyses were confined to just these 11 effect sizes. Applying the moderator variable of experimental design produced three subsets of effect sizes, four associated with studies employing a *pure RI design*, four associated with studies using a *mugshot commitment design*, and three associated with studies using a *transference design.* In the first instance, the metaanalytic Z was not significant, $Z_{ma} = -1.52$, $p < .10$. Mean effect size was $h = -.12$, 95% *CI*: $-.55$ to $+.31$. Thus, the amount of retroactive interference with memory for the target generated simply by interpolation of mugshots does not appear to reduce significantly accuracy (proportion correct) at a subsequent lineup. Identification of a mugshot image as that of the target and the subsequent appearance of that image or person at the lineup is a different matter, however. For the mugshot commitment design, the metaanalytic $Z_{ma} = -4.33$, $p < .0001$, $N_{fs} = 24$, with an associated effect size $h = -.48$ (95% *CI*: $-.24$ to $-.72$). Hence, use of a mugshot commitment design produces a sizeable negative impact upon lineup accuracy, with mean proportion correct of .32 for the mugshot commitment condition versus .55 for the control condition. Apparently identification from mugshots of either a foil or innocent person as the target, interferes sufficiently with a witness's memory representation of the actual target so as to reduce both the hit rate and correct rejection rate for target present and target absent lineups, respectively. Finally, for the transference design, overall proportion correct for the transference condition was .66 and .76 for the control condition, $Z_{ma} = -2.63$, $p < .01$, $N_{fs} = 5$. The corresponding effect size, h, was $-.21$, with a rather broad 95% *CI* of $-.81$ to $+.39$. All in all, the negative impact on proportion correct imposed by a transference design does not appear to be a very trustworthy effect.

A careful analysis of the 11 effect sizes from studies that did not include the target(s) in mugshots showed no discernable relationship between type of experimental design employed and the number of mugshots to which witnesses were exposed, whether a relatively large number or a relatively small one. Hence, with just two different "cuts" of this set of effect sizes, there would appear to be no problem of multicollinearity affecting interpretation of results.

The *number of mugshots* interpolated between exposure to the target and subsequent lineup test was therefore tested as a candidate for moderating the effect of mugshot exposure on eyewitness memory performance. For the 15 tests of the hypothesis that mugshot exposure decreases lineup hit or correct rejection rate, the number of interpolated mugshots ranged from 12 to 600, with a median of 21. In the case of the six studies with above the median numbers of mugshots (41 or more)

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

Mugshot Exposure Effect 32a

and which did not present targets in the mugshots, $Z_{ma} = -1.86$, $p < .05$, $N_{fs} = 2$. The corresponding effect size, h, was $-.14$ (95% *CI*: $-.40$ to $+.12$). Given such a low failsafe N and a confidence interval that includes zero, it is probably best to conclude that a relatively high number of interpolated mugshots is not associated with a deleterious effect on proportion correct at a lineup. Interestingly enough, however, for the five studies with below the median numbers of mugshots and which excluded targets from them, the metaanalytic Z and corresponding effect size were considerably larger, $Z_{ma} = -5.23$, $p < .0001$, $N_{fs} = 46$, $h = -.44$ (95% *CI*: $-.24$ to $-.64$).

Metaanalysis 2: Mugshot Exposure and the False Alarm Rate

All Tests

We first tested the hypothesis that mugshot exposure decreases accuracy at a subsequent lineup by virtue of increasing the false alarm rate. Clearly, unlike Metaanalysis 1, analysis of whether or not the target appeared among the interpolated mugshots was irrelevant. Parallel to our analysis of the proportion correct measure, we subtracted the false alarm rate for the control condition in each instance from that for the experimental condition. Thus, a positive metaanalytic Z would indicate a debilitating effect of mugshot exposure on the false alarm rate, that is to say, by increasing it. The overall false alarm rate at a lineup for mugshot exposure conditions was .37, and the comparable figure for the control conditions was .15, less than half the former rate. The unweighted metaanalytic $Z_{ma} = 9.94$, $p < .0001$, $N_{fs} = 602$. Weighting each of the 17 tests by sample size yielded $Z_{mn} = 9.05$, $p < .0001$, $N_{fs} = 499$. Mean effect size, h, for the unweighted analysis was $+.48$, with a 95% confidence interval extending from $+.26$ to $+.70$. Median effect size was $+.41$. These analyses support the hypothesis that mugshot exposure increases the false alarm rate at a subsequent lineup.

Seeking to determine whether the variability in effect sizes exceeded that expected from sampling error alone, we applied a test for homogeneity of variance to the sample of weighted effect sizes; $\chi^2 = 36.42$, with $df = 16$, $p < .01$. Subsequently, we tested for moderator variables which might account for the substantial variability in effect sizes for individual tests of this hypothesis (effect size range $= -.318$ to $+1.442$, $s = .42$; see Table 1).

Moderator Analyses

Inasmuch as the four studies employing a *pure RI design* only reported overall proportion correct, and false alarm rates for target present lineups could not be inferred due to the false rejection rate not being reported, variability in effect sizes moderated by a pure retroactive interference design could not be ascertained However, four studies using a *mugshot commitment design* did provide tests of whether mugshot exposure in the context of that particular design increases the false alarm rate at a subsequent lineup. There was indeed support for a substantial moderating effect for application of this design, with a mean false alarm rate for the mugshot commitment condition of .53 versus just .20 for the control condition, $Z_{ma} = +6.09$,

This content downloaded from
141fff:ffff:ffff:ffff:ffff on Thu, 01 Jan 1976 12:34:56 UTC
All use subject to https://about.jstor.org/terms

$p < .0001$, $N_{fs} = 51$. The average effect size, h, was $+.72$ (95% *CI*: $+.38$ to $+1.06$). Finally, 11 studies combining mugshot exposure within the context of a *transference design* produced evidence of a significant moderating effect of this particular design on the lineup false alarm rate (.38 for the transference condition as compared with .16 for the control condition). The unweighted metaanalytic Z was $+8.66$, $p < .0001$, and the fail-safe N was 295 additional studies averaging null results. The corresponding average effect size, h, was $+.44$, the 95% confidence interval for which stretched from $+.11$ to $+.77$. It is fair to conclude here that the effect size for the mugshot commitment design was decidedly larger than that for the transference design, some 65% larger.

A careful analysis of the sample of effect sizes again showed no apparent confounding of type of experimental design and the number of mugshots to which witnesses had been exposed. Therefore, we again decided to test for the moderating influence of the relative *number of mugshots* exposed.

In this instance, the number of interpolated mugshots above the median ranged from 16 to 600, and the number below the median ranged from 8 to 15. As occurred with the moderating effect of number of mugshots on proportion correct, exposure of a below-median number of mugshots (eight studies) in this case had a greater impact on the false alarm rate than did exposure to an above-median number of them (nine studies). In the former case, $Z_{ma} = 7.66$, $p < .0001$, $N_{fs} = 166$; mean $h = +.44$ (95% *CI*: $+.26$ to $+.62$). In the latter case, $Z_{ma} = 6.42$, $p < .0001$, $N_{fs} = 128$; mean $h = +.35$ (95% *CI*: $-.18$ to $+.88$). Not only is the effect size for a relatively small number of mugshots exposed some 25% larger than that for a relatively large number of mugshots exposed, but it is a more trustworthy result.

Metaanalysis 3: Transference Errors

All Tests

We began by testing the hypothesis that irrespective of whether mugshots are exposed to eyewitnesses between initial exposure to the target person and subsequent identification test, use of a transference design will increase the rate at which transference errors are made. Here again, parallel to our metaanalysis of false alarm responses generated as a function of mugshot exposure, we subtracted the false alarm rate (to faces previously seen during mugshot exposure) for the control condition in each instance from that for the transference condition. Hence a positive metaanalytic Z would indicate a debilitating effect of transference conditions upon identification accuracy. That is to say, increased likelihood of a transference error decreases the likelihood of a correct identification at a lineup. When a transference condition had been created, the overall rate at which transference errors were made across our sample of 19 independent tests of the transference hypothesis was .33. The comparable figure for the control condition was .16, less than half the rate for the transference condition.

The unweighted metaanalytic Z was 9.24, $p < .0001$, $N_{fs} = 581$ studies. Weighting each of the 19 tests by sample size yielded a metaanalytic Z of 8.89, $p < .0001$, with a fail-safe N of 537 studies. Mean effect size, h, for the unweighted analysis

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

was +.34, with a confidence interval extending from +.08 to +.60. Median effect size was +.35. These analyses support the hypothesis that whether exposed to mugshots or not, presence of a transference condition increases the particular sort of false alarm error known as a transference error.

Applying a test of homogeneity of variance to the sample of weighted effect sizes, $\chi^2 = 64.58$, with $df = 18$, $p < .001$. As was the case with the two previous meta-analyses, the degree of variability exceeded that which would be expected on the basis of sampling error alone. Hence, we sought to determine whether type of transference design and age of the eyewitness could account for some of this variability in effect size.

Moderator Analyses

Considering just the eight transference design studies we have labeled as bystander studies, $Z_{ma} = 4.08$, $p < .0001$, $N_{fs} = 42$. The corresponding mean effect size, h, was +.19 (95% *CI*: −.33 to +.71). For the 11 transference designs which included mugshot exposure, the metaanalytic Z was 8.66, $p < .0001$, with a fail-safe N of 295 studies. In this latter instance, the average effect size, h, was +.44, with a 95% confidence interval which did not include zero, extending from +.11 to +.77. Hence, though statistically reliable in both cases, transference conditions created as a function of witnesses viewing someone at a lineup who had also been seen in a prior mugshot were more potent in producing transference errors (mean transference error rate of .38 versus .16 for the control condition) than were transference conditions created by the viewing at a lineup of a previously seen bystander (transference error rates of .27 and .17, respectively). The former sort of transference design is associated with an effect size more than twice that of the latter.

The second moderator analysis was performed only on the subset of 11 effect sizes associated with transference designs in which witnesses had been exposed to mugshots. This particular sample of effect sizes was chosen, inasmuch as Perfect and Harris (2003) performed separate tests of the transference hypothesis on younger (mean ages of 20–22) and older adults (mean ages of 63–71) in three different experiments, in all of which the witnesses had been exposed to a set of mugshots interpolated between encounter of the targets and subsequent lineup tests. In none of the bystander studies had witness age been employed as an independent variable. Hence, the moderator analysis was performed on a sample of three effect sizes associated with older adult witnesses and eight effect sizes associated with younger adults.

Perhaps not surprisingly, the transference effect was statistically significant for both age groups. For younger adults, $Z_{ma} = 4.61$, $p < .0001$, $N_{fs} = 55$, and for older adults, $Z_{ma} = 9.08$, $p < .0001$, $N_{fs} = 89$. However, older witnesses were much more likely to commit a transference error, with a mean transference error rate of .57 as compared with a rate of .30 for younger adults (see Bartlett & Fulton, 1991); mean effect size, $h = +.85$ (95% *CI*: −.52 to +2.22) for the older adults. The mean effect size for younger witnesses, h, was +.35 (95% *CI*: +.06 to +.64). Given the very small number of tests of the transference hypothesis with older witnesses, and

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

the resulting very large confidence interval, this moderator analysis result should be interpreted with caution.

DISCUSSION

All Tests: Proportion Correct and False Alarms

Even though exposure of eyewitnesses to mugshots may well be a valuable crime investigative tool, we have adduced considerable evidence for the hypothesis that mugshot exposure decreases accuracy for a given witness at a subsequent lineup, both in terms of decreased likelihood of being correct (including hit and correct rejection rates) and increased probability of committing a false alarm. When we excluded those studies which included targets in a set of mugshots as well as at the lineup, we noted that exposure to mugshots decreased proportion correct at the lineup, the mean effect size being $-.28$, with a 95% confidence interval not including zero. The metaanalysis of false alarm rates produced an average effect size of $+.48$, with the lower bound of the confidence interval being $+.26$. However, evidence of the deleterious effect of mugshot exposure on proportion correct at the lineup is on somewhat less solid ground than is the evidence of mugshot exposure elevating the lineup false alarm rate. In the former case, the fail-safe N associated with the unweighted metaanalytic Z requires only 87 additional studies averaging null results to increase the probability of a Type I error to the just significant level of $p = .05$; in the latter case, the fail-safe N is 602 studies. Thus, the effect of mugshot exposure in increasing the false alarm rate is both 70% larger and has greater long-run trustworthiness than its effect in decreasing the proportion correct at subsequent lineups.

Moderator Variables

In addition to establishing the direction and magnitude of the effect of mugshot exposure on proportion correct and the false alarm rate at a subsequent lineup, we had sought to identify variables that might moderate any consistent effects of mugshots on the fidelity of eyewitness memory measurement. In the case of both dependent measures, we identified two different moderators of the effect of mugshot exposure, type of experimental design, and number of mugshots exposed.

The particular method by which mugshot exposure's effects were tested produced strong differences in effect size on both proportion correct and false alarm rate. Mugshot commitment designs yielded the largest effect sizes in both instances, $-.48$ for proportion correct and $+.72$ for the false alarm measure. Yielding smaller deleterious effects on memory performance were transference designs, producing an average effect size of $-.21$ for proportion correct and $+.44$ for false alarms. The moderating effect of pure retroactive interference designs could only be determined for the proportion correct measure, resulting in an effect size of just $-.12$.

It is interesting that mugshot-induced retroactive interference, in its purest form, has only a very limited negative impact upon eyewitness memory. As noted

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

earlier, the pure retroactive interference design would include only previously unseen foils at a lineup. Mugshot commitment and transference designs, on the other hand, both include one or more faces in the ultimate test of memory that had been included in the mugshots. In our sample of studies, the photobiased nature of the lineup in a transference design increased the negative effect size for proportion correct by 75%, from −.12 (effect size for pure RI designs) to −.21. The mugshot commitment design compounds this photobias by including in the lineup not only a person whose face was previously included among the mugshots but also by the previously viewed person having been identified by the witness *as* the target person. If the witness had indeed been willing to identify a previously unseen mugshot as the face of the target person, then clearly the witness must not have had a very distinctive encoding of the target's face, and it would neither be surprising that there would be a lowered probability of a correct identification response nor an increased probability of a false alarm in a subsequent lineup. For our study sample, use of a mugshot commitment design increased the negative effect size for proportion correct by a factor of four, from −.12 to −.48. Use of a mugshot commitment design also increased the effect size for the false alarm rate by approximately 65%, from an effect size of +.44 for transference designs to an effect size of +.72.

Thus, we have been able to provide substantial quantitative support for the empirical generalization that exposure to a mugshot of a suspect increases the probability that the eyewitness will subsequently choose that suspect in a lineup, especially when the real suspect is absent. Not only is the photobiased lineup resulting from the presence of transference conditions productive of such a bias, but identification of a mugshot *as* the suspect (mugshot commitment) substantially furthers the bias. Now Kassin, Tubb, Hosch, and Memon (2001) have reported in the most recent survey of eyewitness testimony experts ($N = 64$), that 79% of the experts believed mugshot-induced bias to be either a "generally reliable" or "very reliable" empirical phenomenon. Similarly, Wise and Safer (2004) in their survey of 160 US judges as to their knowledge and beliefs about eyewitness testimony have indicated that fully 78% of them believed the mugshot-induced-bias phenomenon to be "generally true." Given our results, there would seem little reason that these endorsement rates for both experts and judges should not be even higher in the future. In addition, our results make it clear that the likelihood of mugshot-induced bias is relatively low, unless a person has appeared both in mugshots and at a subsequent lineup.

Interestingly, exposure to relatively low numbers of mugshots had a more negative effect on witness memory than did greater numbers of mugshots. Perhaps it is easier to encode for memory storage a small set of mugshots in a reasonably distinctive fashion than it is for a much larger set. Certainly more distinctively encoded items are more easily retrieved, as they might be later at a lineup test of memory. In addition to a task difficulty explanation of this phenomenon, there is one that is related to motivation or attention: Perhaps exposure to a relatively large number of mugshots motivates witnesses to put minimal effort into encoding the photos, because the police clearly have no real suspect, and the witnesses are just searching. On the other hand, with a relatively small array of photos, witnesses might believe that they are in something like a focused lineup situation in which the police have a suspect; hence, witnesses might well be motivated to pay greater attention to and to

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

engage in more distinctive encoding of the photos, including that of the person who may subsequently appear in a lineup.

Transference Errors

Our final pair of related goals were to determine whether transference designs are reliably productive of increased rates of transference error and if this be the case, whether mugshot exposure is any more productive of transference error than are transference situations not involving mugshot exposure. Answering the latter question will determine whether transference conditions enabled by the criminal justice system, ones including photobiased lineups, result in greater rates of transference error than do transference conditions that are more nearly "accidents of nature." Considered in their entirety, use of transference designs increased the likelihood of transference errors, with an associated effect size of +.34 and a confidence interval not including zero. Transference designs involving mugshot exposure were associated with an effect size (+.44) more than twice that of transference designs having no possibility of a subsequent photobiased lineup (+.19). Perhaps the smaller effect size for innocent bystanders observed at the crime scene obtains for the same reason as the earlier cited false fame effect (Jacoby et al., 1989). The latter effect requires that the stimulus be perceived just well enough to be encoded but not so well encoded that the witness recalls the source. On the other hand, a witness viewing at least a relatively small number of mugshots shortly after the crime likely inspects them with a greater degree of thoroughness, given that one of them may be an image of the perpetrator. Again, the latter more distinctively encoded images are more likely to be mistakenly retrieved at a later target-absent lineup. At any event, both questions we have posed concerning the nature of transference effects can be answered in the affirmative.

Even though Wise and Safer's (2004) survey did not ascertain what their sample of US judges believed to be true with respect to the phenomenon of unconscious transference, Kassin et al. (2001) did find that just 60% of experts believed unconscious transference to be generally reliable or very reliable. We should note that Kassin et al. defined unconscious transference for their sample of experts as, "Eyewitnesses sometimes identify as a culprit someone they have seen in another situation or context" (p. 408). From the present vantage point, this statement is remarkably similar to Kassin et al.'s definition of mugshot-induced bias, "Exposure to mug shots of a suspect increases the likelihood that the witness will later choose that suspect in a lineup" (p. 408). It is not clear whether very many of Kassin et al.'s experts had this similarity in mind when they rated the reliability of the phenomenon of unconscious transference. Given the absence of a relevant metaanalysis, if the experts had been relying only on the number of significant results favoring the transference hypothesis, three out of eight tests in the literature that had been available to them, then the 60% endorsement figure would appear to be unrealistically high. Perhaps, then, some of the experts recognized that mugshot-induced bias is oftentimes just another instance of the transference effect.

Given the results of our metaanalysis of the results of transference studies, the question becomes one of how an expert should now view the reliability of the transference phenomenon. Even the effect size (+.19) associated with the

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

bystander studies also is associated with a metaanalytic Z ($p < .0001$) that is sufficiently reliable that the fail-safe $N = 42$. That is, 42 additional undiscovered studies or studies to be conducted in the future, averaging null results, would be required to increase the probability of a Type I error to the just significant level of $p = .05$. If one were to make a reliability judgment based on just these studies, perhaps an endorsement rate of 60% of experts viewing transference to be a generally reliable phenomenon would not be unreasonable. However, if one properly considers the entire corpus of transference studies as tests of the hypothesis that transference conditions will increase the rate at which transference errors are made, then our results would not make unreasonable an endorsement rate of "generally reliable or very reliable" approaching 100%. After all, the effect size for the entire corpus of transference studies was $+.34$, with a confidence interval not including zero. Furthermore, the metaanalytic Z was associated with a fail-safe N of nearly 600 studies.

Limitations

Two limitations might be proposed to the generality of findings produced by our metaanalyses. First, it might be objected that we should have included as many unpublished studies of the effects of mugshot exposure as we could have uncovered. We shall not review here arguments for including only published studies, arguments made in the method section of this paper. However, as promised, we will present a posthoc analysis demonstrating that even the weakest of our overall metaanalytic findings would not have been changed had we included effect sizes from as many as four additional unpublished studies. For the sake of argument, had we included precisely the same effect sizes as those produced by Steblay's (1992) set of four unpublished studies (15 published effect sizes), our findings regarding the effect of mugshot exposure on proportion correct at a subsequent lineup would have changed very little. Mean effect size, h, would have decreased from $-.15$ to $-.12$, with the confidence interval being almost the same, actually a bit tighter, $-.27$ to $+.03$. Median effect size would have remained the same at $-.17$. The metaanalytic $Z_{ma} = -2.72$, $p < .01$, essentially the same values as we have already reported, and the fail-safe N would have decreased by only one study, from 35 to 34 undiscovered studies.

Second, it might be objected that our relatively small samples of effect sizes might lead either to Type II errors on the one hand or to making too much of statistically significant effects on the other. With regard to the possibility of Type II errors, we would note, for example, that even though a set of four pure RI design studies did not produce a significant moderating effect on proportion correct, a set of four mugshot commitment design studies did so, an effect that was judged to be trustworthy. Thus, small sample sizes, per se, in the present instance did not prevent relatively powerful moderator variables from emerging from our analyses. To guard against making too much of a statistically reliable effect for a moderator variable, we have proposed a reasonably stringent set of minimum criteria for what we have chosen to call a trustworthy effect. In addition, given the small sets of effect sizes that would be involved in any moderator analysis, we have imposed a limit of two moderator analyses for each dependent variable subjected to an overall

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

metaanalysis. In each instance, a second moderator analysis was conducted only after a careful analysis for the possibility of a confounding of the two variables yielded no discernable confounding. This latter step we have taken should have prevented the sort of overinterpretation of results that is based on multicollinearity.

CONCLUSIONS

The results of our three metaanalyses and associated moderator analyses permit us to draw several conclusions concerning the impact of mughot exposure on the fidelity of eyewitness memory. First, prior mugshot exposure decreases accuracy at a subsequent lineup, both in terms of reductions in rates for hits and correct rejections as well as in terms of increases in the rate for false alarms. Second, we have identified two variables, which are statistically reliable moderators of the effect of mugshot exposure, type of exposure design and number of mugshots exposed to the witness, the former being the more important of the two. At least with type of exposure design, there is some possibility of the criminal justice system making some beneficial modifications. Though police investigators do not perform experiments in the scientific sense, they could, for example, reduce the probability of untoward transference effects due to mugshot exposure by the simple expedient of not allowing a person to appear in a lineup whose photograph had been exposed in a prior set of mugshots, but not identified by the witness. Third, we have shown that findings from studies done to try to establish the reliability of the unconscious transference phenomenon and those from studies done to try to establish the reliability of the photobiased lineup are both rightly considered as products of the same basic transference design. In both instances, there is transference of the identity of the target person to an innocent person, a person whose face occurred in proximate temporal context with that of the target. For transference studies included in our metaanalyses, when there were significant increases in transference errors, they were always associated with a proximate temporal context ranging from having the innocent person appear approximately 1 min before the target person to as long as a week afterward.

Clearly there would be benefits both in theory and application to further research efforts being aimed at exploring the boundaries of this zone of temporal proximity to the target person. Future research could also be profitably directed toward determining more precise parameters for the moderating effect of number of mugshots exposed to the witness.

REFERENCES

References marked with an asterisk indicate studies included in the metaanalyses.

Bartlett, J. C., & Fulton, A. (1991). Familiarity and recognition of faces: The factor of age. *Memory & Cognition, 19*, 229–238.

*Brigham, J. C., & Cairns, D. (1988). The effect of mugshot inspections on eyewitness identification accuracy. *Journal of Applied Social Psychology, 18*, 1393–1410.

*Brown, E. L., Deffenbacher, K. A., & Sturgill, W. (1977). Memory for faces and the circumstances of encounter. *Journal of Applied Psychology, 62*, 311–318.

Clark, S. E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior, 29*, 575–604.

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.
*Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987). The reliability of eyewitness identification: The role of system and estimator variables. *Law and Human Behavior, 11*, 233–258.
*Cutler, B. L., Penrod, S. D., & O'Rourke, T. E. (1986). Unconfounding the effects of contextual cues on eyewitness identification accuracy. *Social Behaviour, 1*, 113–134.
Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).
*Davies, G., Shepherd, J., & Ellis, H. (1979). Effects of interpolated mugshot exposure on accuracy of eyewitness identification. *Journal of Applied Psychology, 64*, 232–237.
Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28*, 687–706.
*Deffenbacher, K. A., Carr, T. H., & Leu, J. R. (1981). Memory for words, pictures, and faces: Retroactive interference, forgetting, and reminiscence. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 299–305.
*Dysart, J. E., Lindsay, R. C. L., Hammond, R., & Dupuis, P. (2001). Mugshot exposure prior to lineup identification: Interference, transference, and commitment effects. *Journal of Applied Psychology, 86*, 1280–1284.
*Gorenstein, G. W., & Ellsworth, P. C. (1980). Effect of choosing an incorrect photograph on a later identification by an eyewitness. *Journal of Applied Psychology, 65*, 616–622.
Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56*, 326–338.
Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3–28.
Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67–85.
Kassin, S. M., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the "general acceptance" of eyewitness testimony research: A new survey of the experts. *American Psychologist, 56*, 405–416.
*Lindsay, R. C. L., Nosworthy, G. J., Martin, R., & Martynuck, C. (1994). Using mug shots to find suspects. *Journal of Applied Psychology, 79*, 121–130.
*Loftus, E. F. (1976). Unconscious transference in eyewitness identification. *Law and Psychology Review, 2*, 93–98.
Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*, 252–271.
*Memon, A., Hope, L., Bartlett, J., & Bull, R. (2002). Eyewitness recognition errors: The effects of mugshot viewing and choosing in young and old adults. *Memory & Cognition, 30*, 1219–1227.
*Perfect, T. J., & Harris, L. J. (2003). Adult age differences in unconscious transference: Source confusion or identity blending? *Memory & Cognition, 31*, 570–580.
*Read, J. D., Tollestrup, P., Hammersley, R., McFadzen, E., & Christensen, A. (1990). The unconscious transference effect: Are innocent bystanders ever misidentified? *Applied Cognitive Psychology, 4*, 3–31.
Reisberg, D., & Heuer, F. (2006). The influence of emotion on memory in forensic settings. In M. P. Toglia, J. D. Read, D. F. Ross, & Lindsay, R. C. L. (Eds.), *Handbook of eyewitness psychology: Memory for events* (Vol. 1). Mahwah, NJ: Erlbaum.
Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45*, 1043–1056.
Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*, 183–192.
*Ross, D. F., Ceci, S. J., Dunning, D., & Toglia, M. P. (1994). Unconscious transference and mistaken identity: When a witness misidentifies a familiar but innocent person. *Journal of Applied Psychology, 79*, 918–930.
Shapiro, P. N., & Penrod, S. D. (1986). A meta-analysis of facial identification studies. *Psychological Bulletin, 100*, 139–156.
Simmons v. United States, 390 U.S. 377 (1968).
Steblay, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior, 16*, 413–424.
Williams, G. (1963). *The proof of guilt*. Springfield, IL: Charles C Thomas.
Wise, R. A., & Safer, M. A. (2004). What U.S. judges know and believe about eyewitness testimony. *Applied Cognitive Psychology, 18*, 427–443.

This content downloaded from
141.211.4.224 on Fri, 01 Apr 2022 21:02:13 UTC
All use subject to https://about.jstor.org/terms

RECEIVED by MSC 4/12/2022 1:07:49 PM

APPENDIX C

The Effect of Repeated Lineups

Lin et al. Cognitive Research: Principles and Implications (2019) 4:16
https://doi.org/10.1186/s41235-019-0168-1

Cognitive Research: Principles and Implications

ORIGINAL ARTICLE Open Access

The effects of repeated lineups and delay on eyewitness identification

Wenbo Lin*, Michael J. Strube and Henry L. Roediger III

Abstract

A significant problem in eyewitness identification occurs when witnesses view a suspect in one venue such as a mugshot and then later in a lineup where the suspect is the only previously viewed person. Prior research has documented that the witness may select the suspect from the lineup due either to misplaced familiarity from seeing the mugshot or to their prior commitment from identifying the suspect from the mugshot. Two experiments attempted to minimize these biases by using repeated identical lineups, such that both targets and fillers were repeated, to determine if such a procedure could be useful. Across two experiments, we also varied the delay between seeing the event and the first lineup, as well as the delay between lineups. Despite the use of identical lineups, we continued to observe the effects of commitment and misplaced familiarity, so our procedure did not remove these problems. In addition, we also found that both repeated lineups and increasing delays can influence people's tendency to choose and their willingness to maintain their decisions, regardless of accuracy. Most importantly, however, despite the negative effects of repeated lineups and the relatively long delays used in our experiments, we obtained strong relations between confidence and accuracy when using confidence-accuracy characteristic plots. High confidence responses were associated with high accuracy.

Keywords: Eyewitness identification, Confidence, Decision behaviors

Significance

Eyewitness misidentification in the criminal justice system is a major problem. One contributing factor is the use of repeated identification procedures, whereby witnesses' memories are tested at two separate occasions (e.g., a photo lineup followed by a live lineup) with the suspect as the only person to appear both times. Experimental studies have shown that these procedures introduce a bias so that the suspect, even if innocent, is more likely to be selected in the second identification. The suspect may seem familiar not from having committed the crime, but merely from having been seen previously. Another problem is that witnesses tend to stay committed to their initial identification decision even if it is wrong (the commitment effect).

We developed a procedure to potentially overcome these problems by using identical lineups on two occasions. Because both the suspect and the fillers were the same in the two cases, differences in familiarity should be nullified (although not differences in commitment). However, our procedure still showed effects of misplaced familiarity and commitment. The positive news from the study is that on both identifications and across various delays, witnesses' high confidence identifications were highly accurate. Thus, the confidence-accuracy relation seemed resistant to both repetition and to retention interval. These results add to our growing understanding of eyewitnesses' lineup identification behaviors, as well as further documenting that, on both a first and second test that are not biased, high confidence indicates high accuracy.

Introduction

Erroneous eyewitness identification played a role in over 70% of DNA exoneration cases. One contributor to these misidentification cases involved repeated identification procedures. Witnesses may initially see a mugshot (or a show-up) followed by a lineup, or they may see consecutive lineups (e.g. a photo lineup followed by a live lineup) – and typically the suspect is the only common face between the two. Goldstein (2019) reported

* Correspondence: lin.w@wustl.edu
Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130-4899, USA

Springer Open

© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

about a case in New York City in which a witness identified a man, St. Clair Steward, following the city's standard mugshot procedure. After the crime, detectives had entered broad criteria into a data base: black men between 5′8″ and 6′1″ tall who had been arrested in one of three precincts in New York. The witness leafed through the mugshots until he came to Mr. Steward's, the 31st mugshot, and the witness identified him as the perpetrator. The detectives later placed Mr. Steward in a lineup with other men, and the witness again picked Mr. Steward. He was held in jail for two months, until DNA evidence cleared him, and the case was dropped without fanfare. Of course, without DNA evidence, Mr. Steward may have been tried and convicted. This case is hardly unique. David Lee Wiggins was convicted for the rape of a 14-year-old girl in Fort Worth, Texas. The victim identified him from a photo lineup and live lineups in which Wiggins was the only repeated member in the lineups. He was later exonerated due to DNA evidence. In addition, the famous case of Ronald Cotton (wrongly convicted) and Jennifer Thompson (the victim and witness) also involved repeated lineups in which Cotton was the only repeated person. The perpetrator, Bobby Poole (as later determined by DNA evidence), was in neither lineup, but Thompson identified Cotton in both lineups, although on the first occasion it took her 5 min to do so after narrowing it down to two men.

The common denominator of these cases is that the suspect was the only repeated member in these repeated lineup procedures (single-target-repeated lineups), and studies have shown that such procedures inadvertently bias witnesses toward choosing the suspect. Prior research has focused primarily on repeated identification procedures involving a single repeated member (a guilty or innocent suspect) across multiple identifications (e.g., Godfrey & Clark, 2010; Steblay & Dysart, 2016; Steblay, Tix, & Benson, 2013); all studies showed a bias toward choosing the repeated person on the second occasion. We review the relevant research below, before suggesting a procedure that might mitigate these problems.

A second issue to be considered in our research is that of delay between witnessing an event and being examined on it, as well as the delay between the first and second opportunity to identify the suspect. Of course, longer delays are expected to lead to lower performance and more errors in eyewitness identification. We review the evidence about delay, too, before introducing our experiments that examine both the initial delay between witnessing a crime and a first identification (ID), as well as the subsequent delay between the first and second ID.

The third issue of concern in the present research is measuring the influence of repeated identifications on the confidence-accuracy (CA) relationship as measured by confidence-accuracy characteristic (CAC) plots. Despite problems with repeated identifications, will high confidence responses still indicate high accuracy, as found in virtually all studies in eyewitness memory when assessed by CAC plots on a single unbiased test (Wixted, Mickes, Clark, Gronlund, & Roediger, 2015; Wixted & Wells, 2017)?

Problems with repeated identification procedures

Repeated identification procedures are relatively routine (Behrman & Davey, 2001; Steblay, 2011). On the surface, these procedures may appear beneficial to criminal investigations (e.g., allowing for a confirmation of an eyewitness identification), but laboratory research has consistently shown the negative effects of repeated identifications: mugshots followed by lineup (Brigham & Cairns, 1988; Deffenbacher, Bornstein, & Penrod, 2006; Memon, Hope, Bartlett, & Bull, 2002), prior show-up (seeing a person) followed by a lineup (Godfrey & Clark, 2010; Valentine, Davis, Memon, & Roberts, 2012) and consecutive lineups (Hinz & Pezdek, 2001; Pezdek & Blandon-Gitlin, 2005; Steblay & Dysart, 2016; Steblay et al., 2013). In a show-up usually one person is viewed whereas with mugshots, witnesses often see several or more (31 in the case mentioned above). Unlike in mugshots or show-ups, the suspect in a lineup is presented alongside fillers, which offers some protection against identification errors in that the witness may pick a person known to be innocent or no one at all. Nonetheless, Steblay et al. (2013) still showed that identification errors from the initial lineup were often carried over to the subsequent lineup when the suspect was the only person in both lineups. Similarly, Hinz and Pezdek (2001) found that exposure to an innocent suspect in one target-absent lineup increased the false alarm rate to that innocent person if the same person appeared in a subsequent target-absent lineup. This presumably occurred because only one face was familiar in the second lineup, and this familiarity was misattributed to having seen the person commit the crime rather than from the person appearing in the previous lineup. Thus, the mechanism at work – misattribution of familiarity – may be the same when two successive lineups are viewed with only one person in common, as in the studies when witnesses see a show-up or mugshot and then are given a lineup. In both cases, an innocent person may be selected due to misattribution of familiarity, as occurs in other studies, too (see Jacoby, Kelley, & Dywan, 1989).

Hinz and Pezdek (2001) observed that exposure to an innocent filler in a first target-absent lineup also hurt performance when the second lineup was a target-present lineup (i.e., it included the suspect). The repeated filler (among other fillers) reduced the hit rate of a subsequent target-present lineup (.42) compared to

RECEIVED by MSC 4/12/2022 1:07:49 PM

when the guilty suspect was presented with all new fillers (.76). Together, these studies demonstrated the negative effects of repeated lineups, but their findings reflected procedures where witnesses may have been biased towards choosing the single-repeated target. Our experiments attempt to mitigate this problem by using the same lineup on two occasions, for reasons discussed below.

We have thus far emphasized misattributions of familiarity as a root cause of the problem with repeated IDs with only one person repeated, but at least three different problems potentially exist. Besides misplaced familiarity, two other possible problems are misinterpretation of the police's intention and witnesses' commitment to a prior identification decision. For example, when a particular face is repeated across lineups, witnesses may unintentionally misinterpret the police's intention and believe that the police have identified that particular lineup member as the perpetrator. In addition, witnesses may select an innocent suspect in both the initial and subsequent identification to remain consistent with their initial decision (a commitment effect). The commitment effect has been observed in the misinformation literature (e.g., Schreiber & Sergent, 1998); however, it also applies to the eyewitness identification literature. For example, Valentine et al. (2012) reported that most choosers (88%) in a show-up also selected the same suspect in the subsequent lineup, regardless of whether the initial identification was accurate. This commitment effect has also been observed in other studies (Dysart, Lindsay, Hammond, & Dupuis, 2001; Godfrey & Clark, 2010; Haw, Dickinson, & Meissner, 2007). Eyewitnesses may perceive inconsistent decisions across identifications as an indication of unreliability; therefore, they might feel compelled to stick with their initial decision.

Given these findings, one target for research is whether changes in dual identification procedures could mitigate or eliminate these effects. If both the suspect and fillers are repeated in two lineups, all members would be familiar. Thus, problems due to misplaced familiarity or to misunderstanding what the police (or experimenter) are seeking should, theoretically, be eliminated. Although repeating identical lineups would not necessarily remove the commitment effect, it should yield a better measure of the commitment effect because it would be potentially uncontaminated by the other two issues.

One possible positive use of repeated identical lineups is that performance (identifying the suspect) may improve. Improvement in recall or recognition between two tests is called hypermnesia, and it has been reported in free recall tasks quite regularly (e.g., Erdelyi & Becker, 1974; Roediger & Payne, 1982). Recognition hypermnesia has been reported by Bergstein and Erdelyi (2008), although Payne and Roediger (1987) did not find it with word-list stimuli. If witnesses receive an additional identification opportunity (a subsequent unbiased lineup), would they show hypermnesia? The use of identical lineups may eliminate problems with single-repeated-target lineups and permit greater accuracy on a second test than a first test.

The effects of delay on repeated lineups

The retention interval (or delay) between witnessing an event and being examined on it is a critical variable in all memory research, and eyewitness identification is no exception. However, in repeated identification procedures, there are two relevant delays: the delay between the crime and the initial lineup (the initial delay) and the delay between two repeated identifications (we will refer to this as the subsequent delay).

Numerous researchers have examined how the initial delay affects eyewitness identification, but the effects vary across studies. A longer delay has been shown to decrease the number of correct identifications in some studies (Cutler, Penrod, O'Rourke, & Martens, 1986; Sauer, Brewer, Zweck, & Weber, 2010); however, Egan and colleagues (1977) manipulated retention interval across 2, 21 and 56 days and observed no decrease in correct identifications, but they did observe an increase in false identifications (hence overall accuracy declined). Valentine, Pickering, and Darling (2003) observed a 65% suspect identification rate on an immediate test, but they did not find a reliable decrease in suspect identification rate when the test occurred after 1 or 6 months, although performance was lower (the suspect ID rate fluctuated between 34 and 46%). These null results may have been due to low statistical power. Whereas the effects of delay on identification are not perfectly consistent across studies, they generally suggest poorer identification performance with an increase in delay, confirming over a century of memory research using other procedures (e.g., Nairne, 2008). Importantly, though, studies that have examined the confidence-accuracy relationship with varying delays show that it holds up well across delays as assessed by CAC plots. That is, even though witnesses identify fewer suspects after long delays, if they are highly confident in their identifications, they are generally highly accurate (Palmer, Brewer, Weber, & Nagesh, 2013; Sauer et al., 2010; Wixted, Read, & Lindsay, 2016). However, these studies examined the effects of the retention interval between witnessing a crime and the first identification; no research has examined the effects of the subsequent delay between two lineup identifications on the confidence-accuracy relationship.

As just noted, unlike the initial delay, the effects of the subsequent delay in repeated identifications has received far less attention. The subsequent delay may influence

RECEIVED by MSC 4/12/2022 1:07:49 PM

how eyewitnesses make their second identification decisions and their tendency to switch choices. For example, Godfrey and Clark (2010) found that when the delay between a first and second identification was short (30 min), witnesses generally made consistent decisions (e.g. selecting the same person in both identifications). However, when the delay was one week, they found that witnesses made a combination of consistent decisions and switched decisions (from not picking a suspect the first time to picking one on the second occasion, the no-to-suspect shift). The long delay increased the proportion of no-to-suspect shifts. Steblay et al. (2013) used a 2-week delay between two (non-identical) lineups and they also found a mixture of consistent decisions and no-to-suspect shifts. Although the length of subsequent delay appears to influence witness identification behaviors, no researchers have yet examined CAC plots to see how the confidence-accuracy relationship changes with repeated tests.

Repeated identification and the confidence-accuracy relationship

In addition to reducing suggestibility in repeated lineups and examining the effects of delay, another critical issue is whether the CA relation remains strong across repeated IDs. Eyewitness confidence has been widely used and endorsed as an indicator of accuracy by the U.S. Supreme Court, police, and jurors (Brewer & Burke, 2002; Deffenbacher & Loftus, 1982; Potter & Brewer, 1999). The CA relationship in eyewitness identification has been studied extensively (Bothwell, Deffenbacher, & Brigham, 1987; Juslin, Olsson, & Winman, 1996; Sporer, Penrod, Cutler, & Read, 1995). Through the use of CAC plots, the current state of the evidence indicates that high confidence in an initial lineup indicates high accuracy (Wixted et al., 2015). As noted above, several studies have shown that retention interval does not undermine the CA relationship (Palmer et al., 2013; Sauer et al., 2010; Wixted et al., 2016). However, these studies used single tests, and our research will ask if the relation is maintained on the second lineup when repeated lineups are used.

Steblay and Dysart (2016) recommended against the use of repeated identifications involving the same suspect, but it is unclear from their recommendations whether repeated identifications significantly impair the CA relationship measured using the CAC plots (Wixted & Wells, 2017). There are at least two ways that repeated identifications can affect the CA relation. For instance, witnesses who initially had low confidence in their identification decision may become more confident in a subsequent identification (confidence inflation). Prior research has shown that confidence inflation can occur if witnesses had previously received confirmatory feedback (Bradfield & Mcquiston, 2004; Wells & Bradfield, 1998; Wells, Olson, & Charman, 2003); however, it is uncertain whether repeating testing alone (e.g., repeated identification procedures) causes confidence inflation. Godfrey and Clark (2010, Experiment 2) found that people who consistently chose the suspect were more confident in the subsequent lineup than in the initial show-up, but Steblay et al. (2013) did not find a confidence inflation effect when people consistently chose the suspect in both an initial and a subsequent lineup with only the suspect in common. Alternatively, witnesses may maintain the same level of confidence across lineups (Steblay et al., 2013), perhaps remembering their level of confidence from the first occasion (a confidence carryover effect). Furthermore, the level of confidence expressed at an earlier identification may not only carryover to a subsequent identification, but it may also predict the commitment effect (e.g., high confidence individuals are more likely to commit to their initial decisions than low confidence individuals). In short, confidence inflation could produce poorer confidence and accuracy calibration (e.g., if identification errors are made with higher confidence in Lineup 2 than Lineup 1), but a confidence carryover effect could maintain the CA relationship if it exists in the first lineup. Because no research has examined the effects of repeated lineups (in our experiments, identical lineups) on the CA relationship using CAC plots, our experiments are the first to address this issue.

The present study

The follow questions summarize the objectives of the present study: 1) What are the effects of repeated identical lineups and delay on the tendency to choose from a lineup and on identification accuracy? 2) Given an unbiased subsequent lineup (an identical lineup), are witnesses still committed to their initial decision? 3) Does confidence inflate with repeated lineups? Or does it produce a confidence carryover effect? 4) More importantly, does the confidence-accuracy relationship remain intact despite the effects of repeated lineups and varying delays between them? Because prior studies have shown that both repeated lineups and delay can harm identification accuracy, the primary issue addressed in the current is experiments is whether the combination of these variables significantly impairs the confidence-accuracy relationship.

We report two experiments that investigated the joint effects of repeated (identical) lineups and retention interval on eyewitness identification accuracy and the confidence-accuracy relation. The use of repeated lineups was designed to reduce the problems of selective familiarity of the suspect and the demand characteristic of having only one suspect repeated. Such lineups may be difficult but not impossible to carry out in police departments (e.g. photo lineups could be used on both

RECEIVED by MSC 4/12/2022 1:07:49 PM

occasions, as in our experiments). We also varied the delays between the event and the first lineup (initial delay), as well as between the two lineups (subsequent delay), to determine effects on identification performance and the CA relationship (measured by CAC plots). Manipulating both the initial interval and the subsequent interval permits us, in certain conditions, to examine the effect of repeated testing without a confound of delay (i.e., the second test can be given in one condition at the same retention interval as the initial test in a different condition). In short, repeated lineups and delay tend to co-occur in the real world, and prior research has often only manipulated one of these two variables but not both.

We predicted a greater tendency to choose from the second lineup than the first lineup, both because all faces were repeated (enhancing their general familiarity). That is, we expected to observe nonchooser-to-chooser shifts based on familiarity of the faces, but familiarity would not accrue only to the suspect in target-present lineups, as it has in previous research. On the other hand, we should observe a greater tendency to choose after a short-initial delay than a long-initial delay because people may feel less confident in their ability to make an identification after a long delay due to forgetting. Identification accuracy (the number of guilty suspect IDs) should decrease with an increase in retention interval, and we expect more identification errors in the subsequent lineup than the initial lineup due to misattributions of familiarity accruing to all lineup members. There is no evidence from prior repeated lineups studies about whether the CA relationship will remain intact in a repeated lineup; however, based on prior literature, confidence inflation and confidence carryover effects will probably exist and could potentially harm the CA relationship. The present study also compared confidence judgments and response time in making judgments as indicators of accuracy and found confidence judgments were more predictive of accuracy than response times (see the Additional file 1 for the response time data).

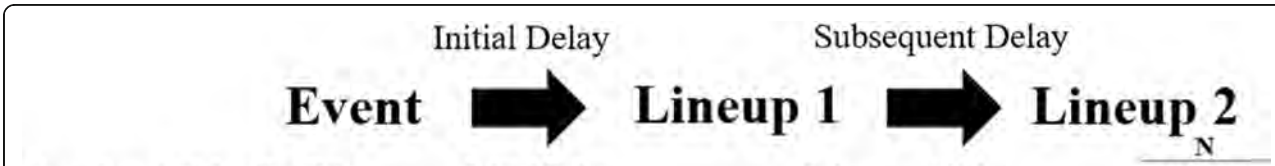
In short, we expected our experiments to confirm some findings from the prior literature and to address several issues for the first time regarding repeated lineups with the same fillers on both occasions. We also have a larger sample size and hence greater numbers of observations and more power than in prior repeated lineup studies.

We will discuss Experiments 1 and 2 jointly because the same stimuli were used, and their overall designs were similar. The main difference between Experiments 1 and 2 was the length of delay between Lineup 1 and Lineup 2 (i.e., the subsequent delay). In Experiment 1, the subsequent delay was always three days. In Experiment 2, the subsequent delay was either one day (a short subsequent delay) or five days (a longer subsequent delay).

Method

Participants

For Experiment 1, 787 participants were recruited via Amazon Mechanical Turk (MTurk), but only 591 participants completed all the sessions. The remaining 196 participants consisted of those who had technical difficulties, completed the follow-up session later than the deadline, or dropped out of the study; their data were not included in the analyses. The number of participants in between-subject conditions ranged from 141 to 154 (see Fig. 1). For Experiment 2, 1158 participants were recruited via MTurk. Only 883 of them completed all the follow-up sessions. The remaining 275 participants whose data were dropped consisted of those who had technical difficulties, completed the follow-up session



Experiment	Delay Condition	Initial Delay	Subsequent Delay	N TP	N TA
1	10min-3days	10 min	3 days	141	143
	3days-3days	3 days	3 days	154	153
2	10min-1day	10 min	1 day	103	109
	10min-5day	10 min	5 days	107	104
	5days-1day	5 days	1 day	111	109
	5days-5days	5 days	5 days	121	119

Fig. 1 The overall design of Experiments 1 and 2. The initial delay determined the length of delay between the Event and Lineup 1, and the subsequent delay determines the length of the delay between Lineup 1 and Lineup 2. All the participants saw the same lineup two times (Lineup 1 and Lineup 2). For each delay condition, the number of participants in target-present (TP) and target-absent (TA) lineup conditions is shown in the right

RECEIVED by MSC 4/12/2022 1:07:49 PM

later than the deadline, or dropped out of the study. The number of participants in between-subject conditions ranged from 103 to 121. All participants were compensated with payment ($2.50).

Design of Experiments 1 and 2

The design of Experiments 1 and 2 were generally similar, as shown in Fig. 1. The procedure always consisted of the event (i.e., the crime video), lineup 1 and lineup 2. In Experiment 1, the initial delay (the delay between the event and lineup 1) was either 10 min or 3 days and the subsequent delay (the delay between lineups) was always 3 days. In Experiment 2, the initial delay was either 10 min or 5 days and the subsequent delay was either 1 day or 5 days. Experiment 1 sought to replicate the effects of initial retention interval on accuracy of choices and on the CA relation. The purpose of the various intervals between lineups (1, 3 or 5 days across the experiments) was to see whether the length of the subsequent delay affected the likelihood of repeated decisions and switched decisions, as well as the CA relationship. Because the subsequent delay was held constant in Experiment 1, we did not expect to find ID differences based on the initial delay conditions (unless ID on initial delay interacts with ID in the subsequent delay). In Experiment 2, we predicted the likelihood of switched decisions to increase as a function of the subsequent delay, but the critical issue is whether the CA relation would also be undermined. We used 1-day and 5-day delays in Experiment 2 to increase the disparity in identification accuracy and (probably) the number of repeated consistent decisions. In sum, there were two delay conditions in Experiment 1 and four delay conditions in Experiment 2 (see Fig. 1). Roughly half of the participants received a target-present lineup and half received a target-absent lineup in both experiments (see Fig. 1 for the exact numbers).

In addition to the eyewitness task, participants in both experiments also completed two distractors tasks (a free recall task and a lexical decision task, described below) after seeing the crime video. Together, the two tasks took approximately 10 min. The purposes were to provide a distractor task and, if needed, to provide covariates for later analyses. However, we did not use them for this latter purpose. In Experiment 1, participants in the 10 min initial delay condition (i.e. the 10 min-3 days condition) saw lineup 1 after the filler tasks. Those in the 3 days–3 days condition did not see lineup 1 until three days later. Similarly, in Experiment 2, participants in the 10 min-1 day and the 10 min-5 days condition saw lineup 1 after the filler tasks. Those in the 5 days-1 day and the 5 days–5 days condition did not see lineup 1 until five days later.

All sessions took place online. Email reminders were sent for each subsequent session along with a link to the online experiment and participant login information. Once they received the email, participants in Experiment 1 were given 36 h to complete the follow-up session and those in Experiment 2 were given 48 h to complete the follow-up session. The longer time window in Experiment 2 was used to minimize the number of late responses.

Materials

The present study used the video and lineup materials from Mickes, Flowe, and Wixted (2012). The video showed a white male (the perpetrator) walking into an unoccupied office and stealing a laptop from the office desk. Viewers had a clear view of the perpetrator as he left the office with the laptop. The video lasted for 23 s and the perpetrator's face was in view for roughly 4–5 s.

Participants received either a target-present or target-absent lineup consisting of six people. The only difference between the two lineups was that the target in the target-present lineup was replaced by a filler in the target-absent lineup. The same lineup was used twice, but the position of the photos was randomized on both occasions. The lineup photos from Mickes et al. (2012) were digitally modified so that all the lineup members wore a black shirt. This step was taken to prevent participants from using the color of clothing as a cue. We also adjusted for the spacing between the top of the lineup member's head and the top edge of the photo. Furthermore, because these photos were taken with different cameras, we made digital adjustments to equate the photo quality. This was to ensure that none of the photos stood out. These changes doubtless made identification in our lineups more difficult than in Mickes et al. (2012). The photos appear in the Additional file 1.

The eyewitness task materials for Experiment 2 were the same as those used in Experiment 1 except another filler in the target-absent lineup was randomly selected to replace one filler in the target-present lineup. In other words, Experiments 1 and 2 had four overlapping fillers in the target-present lineup condition. This was to see whether using another target-replacement filler would produce different frequencies in the overall filler selections.

The categorized words for the free recall task were drawn from the norms of Van Overschelde, Rawson, and Dunlosky (2004). Fifteen words from the *fruit* and the *vegetable* categories were used for the list (Van Overschelde et al., 2004). The lexical decision task consisted of 16 practice trials (4 high- and 4 low-frequency words, and 8 nonwords) and 80 experimental trials (20 high- and 20 low-frequency words, and 40 nonwords). Both the words and nonwords were taken from Yap, Balota, Tse, and

RECEIVED by MSC 4/12/2022 1:07:49 PM

Besner (2008). None of the words from the free recall task overlapped with those in the lexical decision task.

Procedure

After general instructions about the experiment, participants saw a video of the target committing a crime. They were simply instructed to pay attention to the video. Immediately after the video, participants were asked to complete the two distractor tasks: studying and recalling words (about 5 min) and the lexical decision task (another 5 min). In the free recall task, they saw a total of 30 words in a random order for two seconds each, with a 500 ms interstimulus interval. After the last word, participants were asked to recall as many words as possible, in any order, for four minutes. Then the screen automatically advanced to the lexical decision task in which participants were given instructions to decide whether each upcoming letter string represented a word or nonword. There were 16 practice trials followed by 80 experimental trials. Each trial consisted of the following sequence of events: 1) a fixation point (+) presented at the center of the screen for 400 ms, 2) a blank screen for 200 ms. 3) appearance of the stimulus at the fixation point until the participant made a response. Participants were instructed to press the "M" key for words and the "Z" key for nonwords. If they made a correct response, a blank screen was presented for 1600 ms before the fixation point appeared again. For incorrect responses, the word "Incorrect" was shown slightly below the fixation point for 1600 ms. These tasks took at least ten minutes to complete, but the exact duration depended on the participant's pace during the lexical decision task.

In Experiment 1, after the lexical decision task, participants saw lineup 1 in the 10 min-3 days delay condition or were asked to wait for a follow-up email in the next few days (the 3 days–3 days delay condition). Prior to seeing the lineup, participants were given instructions and were told that the suspect may or may not be present in the lineup. They made their identification decision by clicking on the face of the target, and then they clicked the "next" button to submit their choice. If participants judged that the target was not in the lineup, they could reject the lineup via the "not present" button. They were presented with either a target-present or target-absent lineup, and the same lineup was given in lineup 2 (e.g., participants saw the same target-present lineup or target absent lineup for both sessions). The position of the lineup members was randomized in both lineups. Response timing for responding started when participants saw the lineup and ended when they clicked the "Next" button after selecting one of the faces or the "not present" button. Participants were not given any instructions about how fast they should make a decision. After making their identification decision, they were asked to make a confidence judgment using a 0 to 100 scale (0 = Not confident at all; 100 = Very confident). Participants rated their confidence by moving a slider from 0, where the cursor started, to a point on the scale between that value and 100. As they moved the slider, the corresponding numerical value was displayed above the slider. Both lineups used the same procedure, and no feedback was provided about whether participants made the correct choice. At the start of their follow-up session for Lineup 2, they were asked to identify the suspect from the video and were again told that the suspect may or may not be in the lineup. After making their identification, they were asked to provide a confidence judgment using the same scale. At the end of Lineup 2, participants were thanked for their participation and debriefed.

The general lineup identification process was similar for Experiment 2, but unlike in Experiment 1, choosers in Experiment 2 did not have to click on a lineup member's photo and the "Next" button to submit their response. That is, Experiment 1 participants were free to change their decisions before they submitted their final response. In Experiment 2, they were informed that only their first response would be recorded. This was to equate the response time between choosers and nonchoosers, because nonchoosers in both experiments only had to click on the "Not Present" button to submit their response. The confidence judgment section of the experiment was identical in both experiments.

Results and discussion

The results that answer our basic questions are rather numerous and complex. We have divided them into sections that address the issues we intended to address, and we discuss the implications of the results after they are presented.

Repeated identical lineups increased the tendency to choose

Using an identical lineup procedure that made all faces somewhat familiar on the second ID resulted in an increased tendency to choose during the second lineup, just as occurs in other situations. The choosing rate (the percentage of times participants identified a face, whether right or wrong) is shown in Fig. 2 for the conditions of Experiment 1; the corresponding data for all conditions of Experiment 2 is in Fig. 3. The choosing rate for target-present lineups is presented in the top panels, whereas the choosing rate for target-absent lineups is in the bottom panels.

To evaluate our hypotheses, we used the lm4 package in R (Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2017) to conduct a 2 (repeated lineups) × 2 (delay conditions) × 2 (lineup type) generalized mixed model for Experiment 1. Although there were four delay

RECEIVED by MSC 4/12/2022 1:07:49 PM

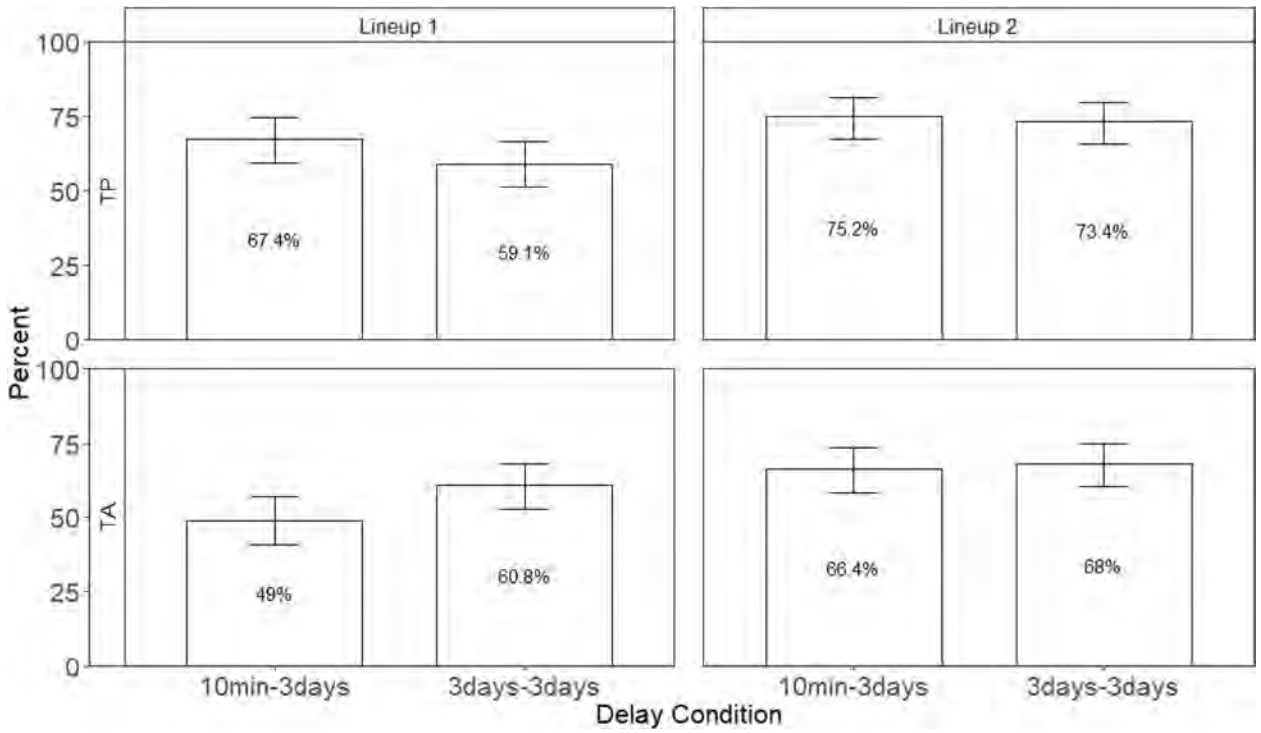


Fig. 2 The percentage of choosers across lineups and delay conditions in Experiment 1. The top panels show the target-present lineups and the bottom panels show the target-absent lineups. Error bars represent 95% confidence intervals

conditions in Experiment 2, two conditions had an initial delay of 10 min and the other two conditions had an initial delay of 5 days. Therefore, we compared the two 10-min conditions with the two 5-days conditions using a 2 (repeated lineups) × 2 (delay conditions) × 2 (lineup type) generalized mixed model. We also conducted additional delay conditions comparisons for Lineups 1 and 2 separately for Experiment 2 (see Additional file 1). The dependent variable was the status of choosing (choosers were coded as 1 and nonchoosers were coded as 0). Each model used the binomial link function. Participants were specified as a random effect for both

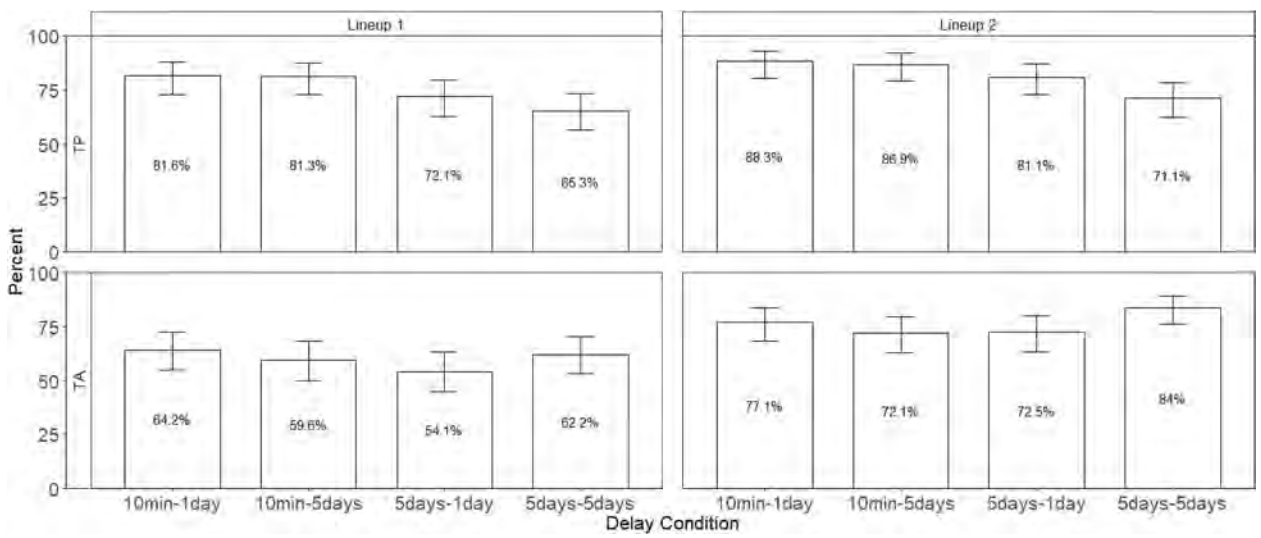


Fig. 3 The percentage of choosers across lineups and delay conditions in Experiment 2. The top panels show the target-present lineups and the bottom panels show the target-absent lineups. Error bars represent 95% confidence intervals

RECEIVED by MSC 4/12/2022 1:07:49 PM

models. Because we were interested in specific comparisons, we constructed a no-intercept model that produced the condition estimates as fixed effect estimates and then we performed specific comparisons (via z-tests) among the condition estimates that corresponded to the effects of interest. These multiple comparisons were conducted using the multcomp package (Hothorn, Bretz, Westfall, & Heiberger, 2012). There is no consensus on the most appropriate method for calculating degrees of freedom in mixed models. Accordingly, many R packages (including multcomp) report estimates divided by their standard errors as z values. For the sample sizes in our studies, this is a reasonable approximation. Adjusted *p*-values (Holm's method) were reported for these multiple comparisons. Unless otherwise indicated, alpha = 0.05.

As predicted, there was a main effect of repeated lineups in both Experiment 1, $z = -4.16$, $p < .001$, and Experiment 2, $z = -5.33$, $p < .001$, showing that regardless of delay, the tendency to choose in Lineup 2 was greater than Lineup 1 (70.7% vs. 59.0% and 79.0% vs 67.4% in Experiments 1 and 2, respectively). In order to observe an increased tendency to choose in the subsequent lineup, most choosers must remain choosers and some nonchoosers must make a positive identification in the second lineup, suggesting that (as discussed below) the commitment and misplaced familiarity effects are likely playing a role.

Although the main effect of delay did not reach significance in Experiment 1, it was significant in Experiment 2, $z = 3.03$, $p < .01$. Participants were generally less inclined to choose after a long-initial delay (5-day delay) than a short-initial delay (10-min delay). This outcome indicates that participants were less confident in their ability to make a positive identification when the initial test was delayed. This interpretation is reflected our Welsch t-test of the higher average Lineup 1 confidence ratings of choosers in the short versus long initial delay in Experiment 1 (61.0% vs 47.3%), t (346.46) = 5.18, $p < .001$, and Experiment 2 (61.4% vs. 40.3%), t (585.74) = 9.94, $p < .001$. Moreover, because there was a main effect of repeated lineups but no interaction between repeated lineups and delay, repeated lineups seem to make participants more inclined to choose regardless of delay. Therefore, even though longer retention intervals made people less inclined to choose and lowered their confidence in their positive identifications, repeated lineups still make witnesses more inclined to choose from a subsequent lineup even with a relatively long delay.

The presence of the guilty suspect in the TP lineup did contribute to the higher TP choosing rate relative to TA choosing rate, although the effect seemed surprisingly small. The main effect of lineup type reached significance in both Experiment 1, $z = -2.78$, $p < .05$, and Experiment 2, $z = -4.77$, $p < .001$. However, there was also a significant interaction between delay and lineup type in Experiment 2, $z = -3.37$, $p < .01$. Follow-up tests showed that choosing rate was greater in TP lineup than TA lineup at the 10-min delay (84.5% vs. 68.3%), $z = -5.39$, $p < .001$, but their choosing rates were no different at a long 5-day delay (72.2% vs. 68.4%). After 5 days, participants were just as likely to choose from a TP as a TA lineup, which seems surprising because retention intervals longer than 5 days are common outside the laboratory. All other effects were not significant, $ps > .05$. The lack of an interaction between repeated lineups and lineup type indicated that the choosing rate for the second lineup increased at about the same level in both TP and TA lineups.

Together, these findings showed that both the effects of retention interval and repeated lineups influence a witness's tendency to choose from a lineup. Participants were less inclined to choose after a long-initial delay than after a short-initial delay and even when they did choose, they were less confident. However, this effect was qualified by the presence of the guilty suspect in the TP lineup. Participants were more inclined to choose from a TP lineup than a TA lineup after a short-initial delay, but they were equally likely to choose from either lineup type after a long-initial delay (Experiment 2). Critically, regardless of delay or the lineup type, repeated lineups increased the tendency to choose. Just as in experiments in which only the suspect is repeated between two lineups, repeating all the faces in two lineups leads to an increase in choosing. We turn next to accuracy.

Repeated identical lineups did not increase accuracy of identifying the suspect

Although participants were more likely to choose someone from a second lineup than a first lineup, they were no more accurate in picking the suspect. Accuracy is represented by the guilty suspect ID rate, which is the number of guilty suspect IDs divided by the total number of target-present lineups. Because the correct decision for a target-present lineup is a guilty suspect identification, all other possible responses in a target-present lineup (e.g., filler identification and "Not Present") are incorrect. The top panels of Figs. 4 and 5 show the percentage of guilty suspect IDs in TP lineups in Experiments 1 and 2, respectively. We conducted a 2 (delay conditions) × 2 (repeated lineups) generalized mixed model for Experiment 1, and a 2 (delay conditions) × 2 (repeated lineups) generalized mixed model for Experiment 2. The dependent variable was accuracy (correct decisions were coded as 1 and incorrect decisions were coded as 0). The analyses here were like those conducted for choosing rates.

RECEIVED by MSC 4/12/2022 1:07:49 PM

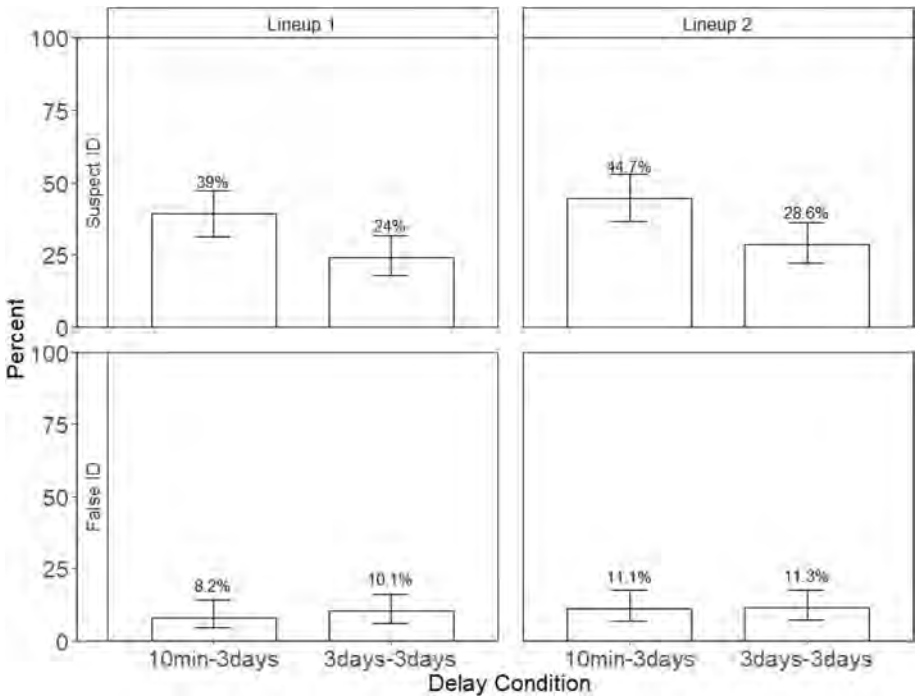


Fig. 4 The top panels show the percentage of suspect IDs and the bottom panels show the percentage of false IDs in Experiment 1. Error bars represent 95% confidence intervals

As predicted, the guilty suspect ID rate decreased as a function of delay in Experiment 1, $z = 3.97$, $p < .001$, and Experiment 2, $z = 6.85$, $p < .001$, but guilty suspect ID rate did not increase (or decrease) with repeated lineups, $ps > .05$ in either experiment. In other words, repeated identical lineups had no effect on overall accuracy. No hypmermnesia – improved recognition with repeated testing as obtained by Bergstein and Erdelyi (2008) in a different paradigm -- occurred. All other effects were not significant, $ps > .05$. In addition, and similar to the effects of delay on choosing rate, the average confidence when picking guilty suspects was greater on a first

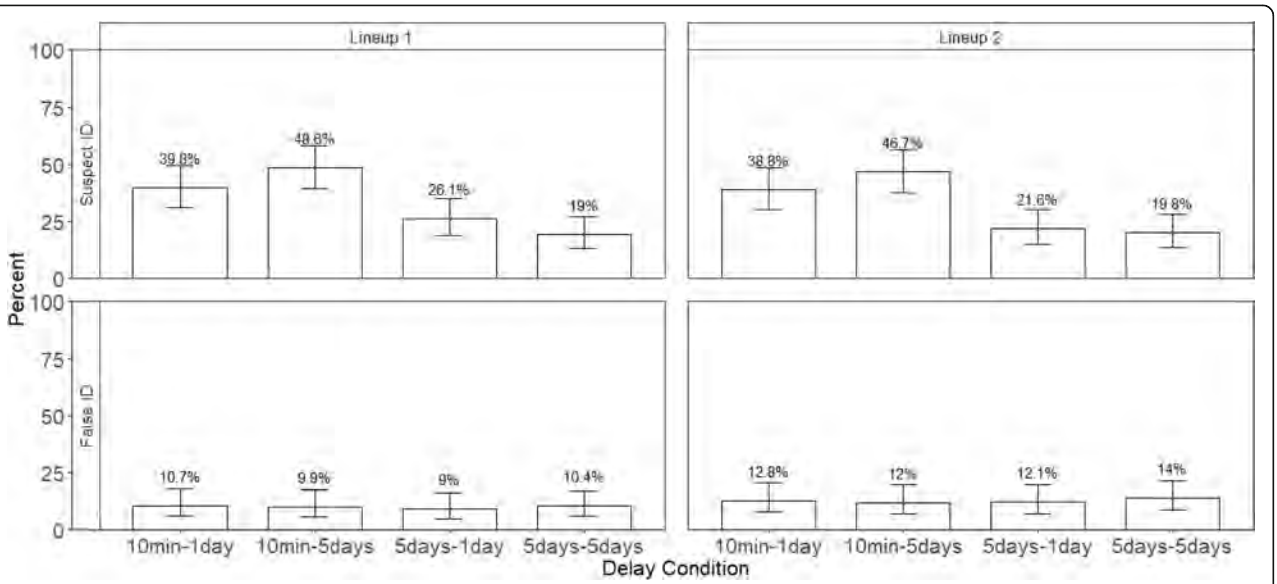


Fig. 5 The top panels show the percentage of suspect IDs and the bottom panels show the percentage of false IDs in Experiment 2. Error bars represent 95% confidence intervals

RECEIVED by MSC 4/12/2022 1:07:49 PM

lineup after a short delay than after a long delay. (Experiment 1: 71.5% vs. 54.6%; Experiment 2: 72.5% vs. 48.4%). Therefore, the effects of delay were reflected in choosing rate, guilty suspect ID rate and the corresponding level of confidence when selecting guilty suspects.

Most choosers repeated their initial decision in the second lineup identification

Although every lineup member was repeated in the second lineup, most identification decisions were carried over from the initial identification to the subsequent identification. The percentage of same and different decisions across Lineups 1 and 2 for Experiments 1 and 2 are shown in Table 1. The "same decision" category consists of identification decisions that were the same across Lineup 1 and Lineup 2 (e.g., selecting the suspect twice, a filler twice or rejecting the lineup twice). The "different decision" category consists of participants making different IDs across the lineups (e.g., selecting the suspect and then choosing a filler, among other possibilities).

Collapsing across all delay conditions, our chi-square goodness of fit test indicated a greater proportion of repeated decisions (consistent decisions) than inconsistent decisions in Experiment 1 (65.0% were consistent decisions), $\chi^2[df = 1] = 53.01$, $p < .001$, and Experiment 2 (60.5% were consistent decisions), $\chi^2[df = 1] = 38.76$, $p < .001$. Further analysis revealed that this pattern was driven primarily by choosers rather than those who rejected the lineup (nonchoosers). A 2 (choosers vs. nonchoosers) × 2 (consistent vs. inconsistent) chi-squared test confirmed this in both Experiment 1, $\chi^2[df = 1] = 11.98$, $p < .001$, and Experiment 2, $\chi^2[df = 1] = 35.66$, $p < .001$. Therefore, choosers were more likely to repeat their decisions compared to nonchoosers. This finding is consistent with choosing rate results reported in the previous section, because most choosers repeated their decisions and some nonchoosers became choosers in the subsequent lineup identification. Thus, we observed an increase in choosing rate in the subsequent lineup.

Table 1 Experiments 1 and 2: Percentages of Same and Different Decisions across Lineup 1 and 2

Experiment	Lineup Type	Delay	Same Decision
1	Target-Present	10 min-3 days ($N = 141$)	70.9
		3 days–3 days ($N = 154$)	64.3
	Target-Absent	10 min-3 days ($N = 143$)	60.8
		3 days–3 days ($N = 153$)	64.1
2	Target-Present	10 min-1 day ($N = 103$)	70.9
		10 min-5 days ($N = 107$)	66.4
		5 days-1 day ($N = 111$)	60.4
		5 days–5 days ($N = 121$)	53.7
	Target-Absent	10 min-1 day ($N = 109$)	63.3
		10 min-5 days ($N = 104$)	54.8
		5 days-1 day (N = 109)	61.5
		5 days–5 days ($N = 119$)	54.6

In an additional analysis, we examined whether TP and TA choosers repeated their decisions in the subsequent lineup to a similar extent. Predictably, our one-proportion z-test showed that most TP choosers remained consistent in Experiment 1 (74.7% were consistent choosers), $\chi^2[df = 1] = 45.5$, $p < .001$, and Experiment 2 (67.6% were consistent choosers), $\chi^2[df = 1] = 40.8$, $p < .001$. A similar pattern occurred with choosers in TA lineups (i.e., incorrect identifications) in both Experiment 1 (66.3% were consistent choosers), $\chi^2[df = 1] = 17.23$, $p < .001$, and Experiment 2 (67.2% were consistent choosers), $\chi^2[df = 1] = 31.2$, $p < .001$. Thus, those who erroneously picked someone in a TA lineup on a first occasion carried over the error to the second lineup nearly two-thirds of the time, confirming a commitment effect. Although identical lineups might be expected to eliminate the bias observed in situations in which a single target is repeated, we still observed a commitment effect for choosers even when all faces were familiar in the second lineups.

Repeated lineups did not lead to confidence inflation, but there was a carryover effect

Repeated lineups could affect the CA relationship in two ways: confidence inflation and confidence carryover effect. We examined whether Lineup 2 identification decisions were made with more confidence than Lineup 1 decisions, especially for repeated decisions (consistent decisions). If Lineup 2 confidence is greater than Lineup 1 confidence, then there is a confidence inflation effect. If Lineups 1 and 2 are similar in confidence, then this outcome may indicate a confidence carryover effect.

Overall confidence in lineups 1 and 2

Aggregating all responses, Lineup 1 identification responses were generally made with more confidence than Lineup 2 responses in Experiment 1, t (590) = 4.78, $p < .001$, and Experiment 2, t (882) = 4.49, $p < .001$. Therefore, witnesses were overall more confident in Lineup 1 than Lineup 2. Next, we examined confidence ratings for consistent choosers and consistent nonchoosers (i.e., witnesses who made the same decisions in Lineups 1 and 2).

Consistent choosers and no-to-suspect choosers

No inflation in confidence ratings occurred for choosers who repeated their initial decisions in the second lineup (see Table 2). Consistent choosers were equally confident across lineups in Experiment 1 (56.0% to 57.3%) and Experiment 2 (54.5% to 53.0%), with Bayes factors (BF_{01}) of 8.09 and 4, respectively, indicating moderate support for the null hypothesis. However, they were more confident

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 2 Experiments 1 and 2: The Average Confidence Ratings for Consistent Choosers and Nonchoosers

Experiment	Choosing Status	Lineup 1	Lineup 2
1	Choosers ($N = 247$)	56.0	57.3
	Nonchoosers ($N = 137$)	62.4	53.1
2	Choosers ($N = 401$)	54.5	53.0
	Nonchoosers ($N = 133$)	57.0	48.1

than choosers who initially rejected but chose someone at the second lineup (no-to-suspect shift). In Experiment 1, our independent t-test indicated that witnesses who made no-to-suspect shifts in Lineup 2 were less confident than consistent choosers in Lineup 1(44.4% vs. 56.0%), t (350) = 3.84, $p < .001$, and at Lineup 2 (44.4% vs. 57.3%), t (350) = 4.14, $p < .001$. In Experiment 2, our independent t-test revealed that witnesses who made no-to-suspect shifts in Lineup 2 were again less confident than consistent choosers at Lineup 1 (43.0% vs. 54.5%), t (554) = 4.41, $p < .001$. Due to unequal variance in Lineup 2, we conducted a Welsch t-test and obtained the same pattern of results for our comparison of no-to-suspect shifts and consistent choosers at Lineup 2 (43.0% vs. 53.0%), t (309) = 4.02, $p < .001$. Therefore, although there was no overall confidence inflation for consistent choosers, they were more confident than those who initially rejected but chose someone only in the second lineup.

Consistent correct and incorrect IDs

While consistent choosers were equally confident in both lineups, the level of confidence expressed did differ across response types (e.g., guilty suspect IDs vs incorrect IDs). For example, collapsing across delay conditions for Lineup 1, average confidence from consistent guilty suspect IDs in TP lineups was greater than consistent incorrect chooser IDs in TA lineups (68.0% vs. 47.6% and 66.6% vs. 51.5% in Experiments 1 and 2, respectively). The same pattern held for Lineup 2 (66.0% vs. 51.5% and 65.8% vs. 50.4%). Thus, consistent correct choosers (those who made guilty suspect IDs) were generally more confident than those who made consistent incorrect IDs in TA lineups.

Consistent nonchoosers

On the other hand, our paired t-test indicated that nonchoosers who repeated their decisions in the second lineup actually became significantly less confident in both Experiment 1 (62.4% to 53.1%), t (136) = 4.33, $p < .001$, and Experiment 2 (57.0% to 48.1%), t (132) = 4.62, $p < .001$. Therefore, without post-ID feedback on the first lineup, a second identical lineup did not cause confidence inflation for consistent choosers, but it made consistent nonchoosers less confident about their decisions. One possible interpretation is that the familiarity of all members of the lineup on the second occasion made it more difficult for participants to reject the lineup and thus, even when they did so, they expressed less confidence in their decision.

Correlation between lineups 1 and 2

Although there was no confidence inflation, Lineup 1 confidence was positively correlated with Lineup 2 confidence. In Experiment 1, we obtained a high correlation for both consistent choosers, r (245) = .77, $p < .001$, and for consistent nonchoosers, r (135) = .63, $p < .001$. In Experiment 2, we similarly found a high correlation for consistent choosers, r (399) = .82, $p < .001$, and for consistent nonchoosers, r (131) = .69, $p < .001$. Together with the average confidence results, consistent choosers and nonchoosers did not show an overall confidence inflation effect; however, their initial confidence tended to be correlated their subsequent identification confidence, suggesting the presence of a confidence carryover effect.

Initial confidence predicted the commitment effect

Because most participants repeated their initial decisions and their initial confidence was correlated with their subsequent identification confidence, we examined whether the level of confidence expressed at Lineup 1 was predictive of repeated decisions at Lineup 2. Said differently, were witnesses who were highly confident in their first decision (correct or incorrect) more likely to provide the same decision on the second lineup than those who were less confident in their initial decisions? To evaluate this hypothesis, we categorized participants by whether they made a consistent or inconsistent decision across the two lineups. Then using the level of confidence participants expressed during Lineup 1, we calculated the proportion of these participants who repeated their decisions in Lineup 2. Specifically, we were interested in the proportion of consistent and inconsistent decisions at a high confidence cutoff (90% or above). Table 3 shows the proportion of consistent responses above and below a confidence rating of 90% in both experiments.

Aggregated across Experiments 1 and 2, our 2 (above vs. below 90% confidence) × 2 (consistent vs. inconsistent responses) chi-square test showed that the proportion of consistent responses above 90% confidence was greater than the proportion of consistent responses below the cutoff (76.3% vs. 60.6%, respectively), $\chi^2[df = 1] = 14.9$, $p < .001$. In addition. we decomposed overall results into data for choosers (middle columns of Table 2) and nonchoosers (right columns of Table 2). The greater consistency of among participants who rated their confidence at 90% or more held for both sets of participants, but our 2 (chooser vs. nonchooser) × 2 (consistent vs. inconsistent responses) chi-square test

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 3 Experiments 1 and 2: The Percentages of Consistent Responses Above and Below a Confidence Rating of 90%

Experiment	Delay	Overall		Chooser Only		Nonchooser Only	
		Above	Below	Above	Below	Above	Below
1	10 min-3 days	79.0	63.8	88.2	70.3	71.4	54.1
	3 days–3 days	73.0	63.0	81.8	68.8	69.2	52.6
	Overall	76	63.3	85.7	69.5	70.2	53.3
2	10 min-1 day	80.0	64.8	87.0	73.3	57.1	43.1
	10 min-5 days	81.5	57.6	88.9	64.1	66.7	41.5
	5 days-1 day	63.0	60.8	100	65.7	45.5	51.4
	5 days–5 days	75.0	53.1	100	58.8	57.1	42.5
	Overall	76.5	58.8	90.2	65.3	55.9	44.9

indicated that the effect was greater for choosers (88.6%) than for nonchoosers (64.2%), $\chi^2[df = 1] = 13.2$, $p < .001$.

In short, participants expressing high confidence were more likely to make the same choice a second time than were those expressing lower confidence on the first lineup. Thus, confidence is not only associated with identification accuracy, but initial confidence can be a predictor of the commitment effect (whether witnesses would repeat their initial decision in a subsequent lineup). Putting this outcome together with our previous findings, consistent witnesses generally maintained similar levels of confidence across repeated lineups, but their initial confidence predicted the likelihood that they would provide the same decision (right or wrong) on a subsequent identification. We also considered the effects of subsequent delay on consistent decisions; while the results were in the predicted direction with less consistent decisions after long delays, most of the analyses did not reach significance (see Additional file 1).

A strong confidence-accuracy relationship was obtained in all conditions

The critical question for the criminal justice system is how repeated lineups and delays affect the confidence-accuracy relationship. Briefly, our results show a strong CA relationship even with repeated (identical) lineups and even with long delays. Repeated lineups did not harm the confidence-accuracy relationship in our experiments, but the procedure did not improve it either.

Figures 6 and 7 show the confidence-accuracy characteristic (CAC) plots for Lineups 1 and 2 in Experiments

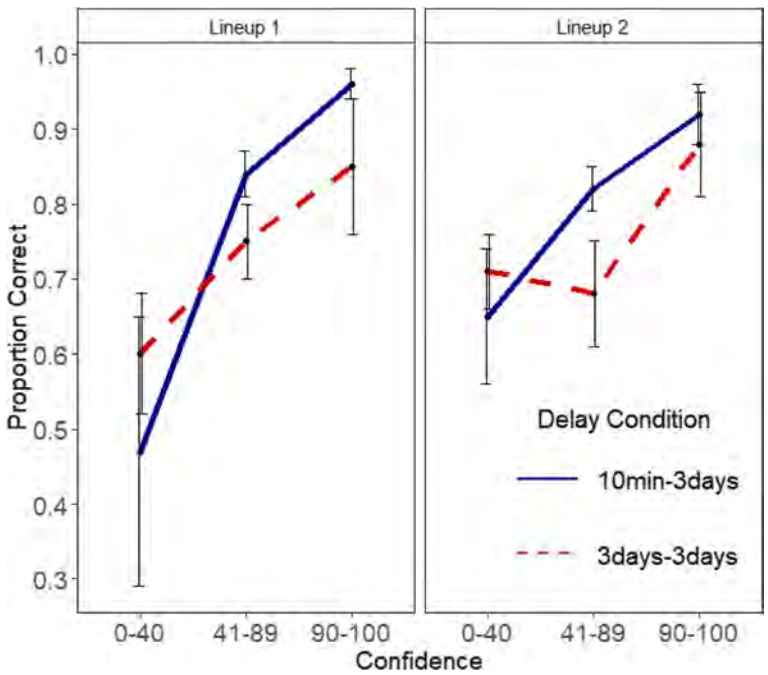


Fig. 6 Experiment 1 CAC plots for Lineups 1 and 2. Error bars are standard errors

RECEIVED by MSC 4/12/2022 1:07:49 PM

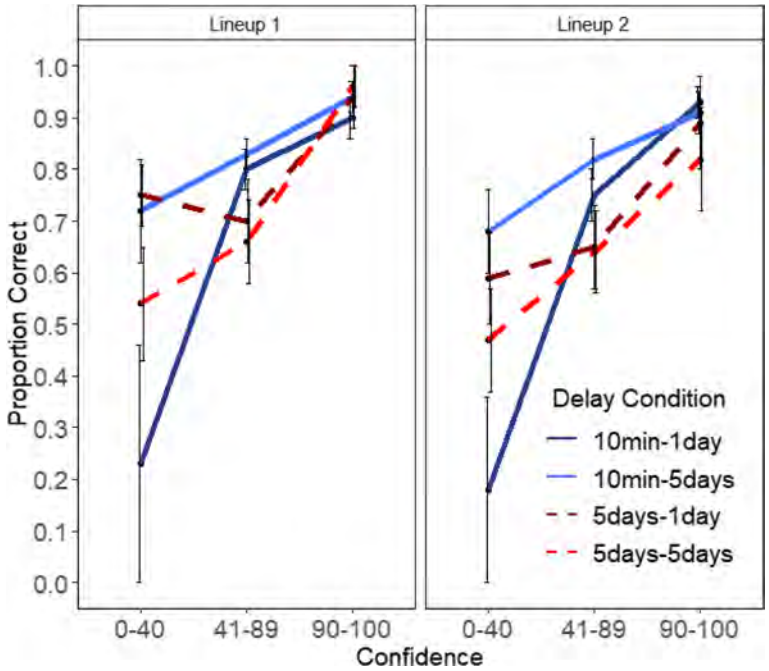


Fig. 7 Experiment 2 CAC plots for Lineups 1 and 2. Error bars are standard errors

1 and 2, respectively. Following the procedure from Mickes (2015) and Wixted et al. (2015) for constructing CAC plots, accuracy was calculated for each confidence bin as the number of correct suspect IDs / (correct suspects IDs + false IDs). Because there was no designated innocent suspect in the target-absent lineups, the sum of false IDs for a particular confidence bin (e.g., all the false identifications made with a confidence rating of 0–40%) was divided by 6, the lineup size (Mickes, 2015). The Additional file 1 contains the distribution of suspect and false IDs across three confidence bins (Low: 0–40%, Medium: 41–89%, and High: 90–100%). The standard error bars in the figures were computed using a jackknife procedure.

We observed a positive relationship between confidence and accuracy in both experiments (Figs. 6 and 7) even in the repeated lineup conditions and even with long delays. In Experiment 1, accuracy in the highest confidence bin for the first lineup (after 10 min) was .96 and it was .92 on the second lineup given 3 days later. When the first lineup was given after 3 days, accuracy was .86 and it was .88 on the repeated lineup 3 days later (see Fig. 6). Comparable Lineup 1 accuracy levels in Experiment 2 for the 10 min-1 day, 10 min-5 days, 5 days-1 day, and the 5 days–5 days were .90, .94, .95, and .96, respectively. The corresponding Lineup 2 accuracy levels were .93, .91, .90, and .83 (see Fig. 7). Because the error bars overlapped for the 90–100% confidence bin, there was no difference in accuracy in the various delay conditions and across the two lineups.

Consistent with prior literature (Palmer et al., 2013; Sauer et al., 2010; Wixted et al., 2016), high confidence continued to be associated with high accuracy even after a long initial delay and/or after a long subsequent delay. In addition, our study shows that this strongly positive relationship also holds true for repeated lineups. This finding, however, is not surprising considering the results in the previous sections. Most participants repeated their initial decision in the subsequent identification. Those who made a high confidence response for the first lineup were generally more likely to do so again for the second lineup. Yet despite making the same decision, they did not increase their confidence (i.e., they showed a carryover of confidence rather than its inflation). Thus, there were no significant differences in accuracy between Lineup 1 and Lineup 2 at the 90–100% confidence bin. We do not discuss confidence effects in the other bins, because only initial high confidence decisions are critical in eyewitness situations.

General discussion

Prior research with repeated lineups (or show-ups or mugshots prior to lineups) has shown that the procedure biases selection toward the suspect, the only face repeated on the two occasions. We tried to overcome this problem by using identical lineups so that all faces were repeated, albeit in a different random order. Despite this change in procedure, our results were similar to those of prior studies that used a single repeated target (Hinz & Pezdek, 2001; Pezdek & Blandon-Gitlin, 2005; Steblay et

RECEIVED by MSC 4/12/2022 1:07:49 PM

al., 2013). When the entire lineup was repeated, and regardless of whether the suspect was in the lineup, witnesses were generally more willing to choose from a second lineup than from the initial lineup. In addition, this increase in choosing did not improve accuracy of responding. No recognition hypermnesia was obtained, as found previously (Bergstein & Erdelyi, 2008; although see Payne & Roediger, 1987). Thus, our revised procedure that eliminated the most basic flaw in all previous repeated identification procedures – only one repeated face – still did not overcome the main problems already identified in repeated lineups – familiarity and commitment effects. Besides replicating prior findings with our new procedure, we also confirmed that witnesses were less able to identify the guilty suspect after a long delay than a short delay; however, they were also less willing to make a positive identification, demonstrating their awareness of their declining retention. This interpretation is further supported by the lower confidence ratings for positive identifications after a long delay than a short delay. In addition to these findings that largely replicate past work, our results extend our understanding of repeated lineup procedures in four basic ways. We discuss these in turn and relate each finding to the prior literature.

First, we showed that, regardless of the retention interval and whether or not the lineup includes the suspect, repeated lineups exert similar effects on witnesses' willingness to make a positive identification. When a lineup was repeated, witnesses were willing to make a positive identification despite their initial reluctance, as we observed in witnesses switching from no choice on a first lineup to choosing a suspect on the second lineup. This outcome replicated across our various initial- and subsequent-delay conditions and it is likely due to general increased familiarity of faces on the second occasion. This tendency to select a face based on familiarity counteracts the trend noted above that, on a first test, witnesses are reluctant to select a face the longer the retention interval and even when they do so, they are less confident in their choice. However, according to Godfrey and Clark (2010), repeated identification procedures may also heighten witnesses' expectations that the suspect would be present in the lineup. Witnesses in the present study may increase choosing on the second lineup because of a shift in their decision criterion or due to misplaced familiarity, but their decision to choose was not simply due to a single-repeated target because we used identical lineups.

Second, we found that witnesses generally remain committed to their initial decision when they pick a suspect from on a repeated lineup, even when they were not biased towards choosing a single repeated target. Like studies that used a single repeated target (Memon et al., 2002; Steblay et al., 2013), we also observed the commitment and misplaced familiarity effects with identical lineups. Because all six faces were repeated in our identical lineup procedure, our results reflected the commitment and misplaced familiarity effects that could not be due to misinterpretation of the police's (experimenter's) intention or misplaced familiarity from a single-repeated target. Therefore, we can conclude that the misplaced familiarity and commitment effects are inherent in repeated lineup procedures even without the problem of single-face repetition. These results suggest that source monitoring (Johnson, Hashtroudi, & Lindsay, 1993) is an issue in both repeated lineups with a single-repeated target and identical lineups. When witnesses are shown a second lineup, they may misattribute their memory of faces from the first lineup for their memory of the original event. As a result, some witnesses may remain committed to their initial decision and others may choose to make a positive identification only in the second lineup. Such source monitoring errors seem ubiquitous in eyewitness research, both in eyewitness identification and from misinformation provided after witnessing a crime (Lindsay & Johnson, 1989). Our results can also be interpreted by Godfrey and Clark (2010)'s decision criterion explanation. Witnesses who made a no-to-suspect shift at Lineup 2 were less confident than consistent choosers at Lineups 1 and 2. If we consider the average confidence expressed for positive identifications as not only a measure of witness certainty but also as an indicator of the willingness to make a positive identification based that level of certainty (i.e. decision criterion), then a lower average confidence in one group relative to another can suggest differences in decision criterion (the acceptable level of certainty to make a positive identification). It is possible that both problems in source monitoring and changing decision criteria can influence the decision process involved in repeated identifications.

Third, we showed that the initial confidence rating is a good predictor of the commitment effect. Consistent with Koriat (2012)'s idea that confidence is a predictor of repeated judgments, we also showed that high confidence identification decisions (regardless of accuracy) are more likely to be repeated in a subsequent identification than low confidence decisions. In addition to causing the same decision, the initial level of confidence is often carried over to confidence in selecting a suspect in the subsequent lineup. Similar to Steblay et al. (2013), the consistent choosers in the present study were also equally confident across in both lineups. However, consistent nonchoosers did become less confident, which might suggest that prior familiarity can make it hard for witnesses to confidently reject a later lineup, as in other studies of misplaced familiarity (e.g., Jacoby et al., 1989). Nevertheless, both confident choosers and nonchoosers often repeat their initial decision in a subsequent identification.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Fourth, and perhaps most importantly, our experiments showed that a strong confidence-accuracy relationship remains true for repeated lineups, even ones occurring after 6-day and 10-day retention intervals since viewing the crime. Prior repeated lineup studies that reported the commitment and misplaced familiarity effects (Steblay & Dysart, 2016; Steblay et al., 2013) did not examine the influence of these effects on the confidence-accuracy relation. Our experiments lead us to conclude that the confidence-accuracy relationship remains strong despite these biasing effects. In addition, the strong relation held for all initial- and subsequent-delay conditions. Because high confidence decisions are often carried over to a subsequent lineup and the level of confidence is also often maintained (rather than inflated), repeated lineups did not significantly impair the confidence-accuracy relationship. Thus, confidence is not only predictive of accuracy across various retention interval (Palmer et al., 2013; Sauer et al., 2010), but it remains a good indicator of accuracy even for repeated lineups.

One surprise in our data is that no confidence inflation occurred on the second lineup of our repeated lineups. One great problem in the criminal justice system has been the growth in confidence over time, such that a low confidence initial judgment becomes a high confidence judgment in the courtroom. This problem is one of the root causes of people being convicted by mistaken eyewitness identification. In *Convicting the Innocent*, Brandon Garrett (2011) wrote about people convicted of crimes and then later exonerated by DNA evidence, "I expected to read that these eyewitnesses were certain at trial that they had identified the right person. They were. I did not expect, however, to read testimony by witnesses at trial indicating that they had earlier had trouble identifying the defendants … Yet in 57% of these trial transcripts (92 of 161 cases), the witnesses reported that they had *not* been certain at the time of their earlier identifications" (p. 49). He identified several causes of confidence inflation over time, such as confirmatory feedback from the police and the fact that people are often repeatedly tested. Yet in our experiments, repeated testing did not lead to confidence inflation. Our experiments are the first to use identical repeated lineups, but our results are similar to Steblay et al. (2013)'s repeated lineups with a single-repeated target. Steblay and colleagues also did not observe a confidence inflation effect. On the other hand, Godfrey and Clark (20102), Experiment found that people who consistently chose the suspect were more confident in the subsequent lineup than in the initial show-up. Therefore, confidence inflation may be possible in some circumstances, but further research is required to determine the cause. In addition, research participants and actual witnesses might approach repeated identifications differently. Compared to research participants, actual witnesses may repeatedly reenact the crime mentally over time, imagining the scene vividly and if they remember it with the wrong perpetrator as the assailant in the crime, their confidence in their judgment is likely to grow. Future research should examine the effects of repeated testing on and imagination inflation on actual witnesses.

In summary, despite our attempts to remove the biases associated with a single-repeated target, our identical lineup procedures still showed the negative effects previously reported in other repeated lineup studies. The good news is that these effects did not significantly undermine the confidence-accuracy relationship, at least in our studies with identical lineups. Therefore, when repeated lineups are not avoidable, confidence can still be a useful indicator of accuracy. Of course, this outcome needs to be replicated in the other kinds of single-target repeated lineups now being used.

Limitations

We acknowledge there are some limitations with the present study. For example, we only used one set of lineup materials. Different eyewitness identification materials can vary in the level of difficulty, which can potentially produce different results in performance measures (e.g., choosing rate and guilty suspect ID rate). Our lineup, due to the changes we made in shirt color, was quite difficult. On the positive side, we show that even with a difficult lineup, the confidence-accuracy relation is strong: High confidence still indicates high accuracy. Thus, in this sense, the difficulty of our single lineup makes the strong CA relationship more convincing.

Another issue is that our experiments did not randomly select lineup members from a pool of faces (a filler pool), which limits the generalization of our results across different arrays of fillers. However, some of our findings, such as the commitment effect, were found in studies that had used materials different from ours. Nevertheless, we of course recommend that the effects reported in the present study be replicated using other sets of eyewitness materials.

Implications

Repeated lineup procedures

Steblay and Dysart (2016) had recommended against the use of repeated identification procedures involving a single repeated target; however, it was unclear from their recommendations whether repeated lineups significantly impair the confidence-accuracy relationship and whether identical lineups would be a viable alternative. The present study addressed these two issues. We showed that identical lineups still displayed the negative effects previously reported in other repeated identification

RECEIVED by MSC 4/12/2022 1:07:49 PM

studies involving a single-repeated target (Godfrey & Clark, 2010; Hinz & Pezdek, 2001; Pezdek & Blandon-Gitlin, 2005; Steblay et al., 2013). However, we also showed that the confidence-accuracy relationship was not significantly impaired by such procedure. This remained true across all of our delay manipulations. Of course, it remains to be seen if this strong relationship holds when only one face (the suspect's) is repeated in two lineups or from a mugshot (or show up to a lineup).

Although repeated lineups did not harm the confidence-accuracy relationship, we do not argue for the general use of repeated lineups, because they have other costs. Repeated lineups led to an increased tendency to choose a suspect, regardless of delay. With only a single lineup, the greater the retention interval between witnessing the crime and seeing the lineup, the less likely witnesses are to choose a suspect and the lower their overall confidence when they do choose. However, when a lineup was repeated, people were willing to choose and this choosing sometimes occurred on the second lineup for witnesses who had declined to choose on the first lineup. Critically, repeated lineups did not improve the suspect ID rate or the confidence-accuracy relationship.

We agree with the overall recommendations from Steblay and Dysart (2016) that repeated lineups are best avoided, but it is unlikely that police will completely abandon all forms of repeated identification procedures. Thus, extra precautions should be taken to administer these procedures. As shown in the present study, using repeated lineups with the same fillers to remove the typical biases found in repeated identification procedures with a single face repeated were not enough to eliminate the negative effects. The encouraging news is that these negative effects were enough not to undermine the confidence-accuracy relationship. Therefore, confidence is still a useful indicator of witness accuracy when repeated identification is unavoidable, assuming this finding generalizes to the single repeated-lineup situation.

Identification behaviors

Research on eyewitness identification often focuses on accurate performance, which is of course critical. Our research examined other behaviors that often receive little attention, such as the witnesses' tendency to choose (whether right or wrong). Our experiments showed that both the length of initial delay and repeated lineups can affect the rate of choosing. People were more likely to choose after a short delay than after a long delay on a first test, and they were more likely to correctly identify the suspects and were more confident when doing so on the initial test. In addition, those who made a positive identification with high confidence were more likely to repeat their decisions and maintain their level of confidence in a subsequent lineup. Overall, witnesses were likely to choose on a second lineup than on a first lineup.

Witness confidence

Our results add to the growing body of evidence that confidence of the witness is a powerful indicator of accuracy; the novel feature of our research is that we showed a strong confidence-accuracy relation on a second lineup as well as a first lineup (as in prior single lineup studies; Wixted et al., 2015; Wixted & Wells, 2017). In addition, we showed that this strong CA relationship exists even after a long delay (ten days) since the original viewing of the crime scenario. Of course, this conclusion may be limited to our conditions of having the same filler faces in the two lineups, a procedure not used in prior studies. Moreover, we have also shown that the confidence in the initial lineup decision also predicts the commitment effect (whether the same decision would be made when viewing the second lineup). In other words, not only does confidence predict accuracy, it also predicts the reproducibility of an identification decision.

Conclusions

In summary, the effects of commitment and misplaced familiarity are present in both repeated lineups involving a single-repeated target and identical lineups, suggesting that these negative effects are inherent in repeated lineup procedures and were not limited to repeated identification procedures involving single-face repetition. In addition, we showed that across two experiments, repeated lineups and increasing delays contribute to people's tendency to make a positive identification and their willingness to maintain their decisions, regardless of accuracy. Critically, however, the confidence-accuracy relationship was not significantly impaired by the negative effects of repeated lineups and delay. High confidence continued to be associated with high accuracy.

Additional file

Additional file 1: The supplemental materials provide additional analyses of the results from our two experiments. (DOCX 2102 kb)

Abbreviations

BCa: bias-corrected and accelerated; CA: confidence-accuracy; CAC: confidence-accuracy characteristic; RT: response time or reaction time; TA: target-absent; TP: target-present

Acknowledgements

This research is based on a master's thesis submitted by the first author under the supervision of the third author. We thank Eylul Tekin, Oyku Uner, and Dave Balota for their helpful comments and John Wixted for providing the eyewitness materials.

Authors' contributions

WL and HR conceived the idea for these experiments. WL programmed the experiments, collected data, and analyzed the results. MS assisted with the

RECEIVED by MSC 4/12/2022 1:07:49 PM

analyses and provided suggestions. WL and HR were the primary authors on the report. All three authors approved the final version.

Funding
The research was funded by Research Fund 94070B to HLR from Washington University.

Availability of data and materials
The datasets and supplementary analyses supporting the conclusions of this article are available in the Open Science Framework repository, https://osf.io/bc3zu/.

Ethics approval and consent to participate
The study was approved by Washington University Institutional Review Board. All the subjects received informed consent at the beginning of the study and they were debriefed at the end of the study.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 12 July 2018 Accepted: 3 May 2019
Published online: 13 June 2019

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using {lme4}. *Journal of Statistical Software*. https://doi.org/10.18637/jss.v067.i01.

Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior*, *25*(5), 475–491.

Bergstein, J., & Erdelyi, M. (2008). Recognition hypermnesia: How to get it. *Memory*, *16*(7), 689–702. https://doi.org/10.1080/09658210802169095.

Bothwell, R. K., Deffenbacher, K. A., & Brigham, J. C. (1987). Correlation of eyewitness accuracy and confidence: Optimality hypothesis revisited. *Journal of Applied Psychology*, *72*(4), 691–695. https://doi.org/10.1037/0021-9010.72.4.691.

Bradfield, A., & McQuiston, D. E. (2004). When does evidence of eyewitness confidence inflation affect judgments in a criminal trial? *Law and Human Behavior*, *28*(4), 369–387. https://doi.org/10.1023/B:LAHU.0000039331.54147.ff.

Brewer, N., & Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and Human Behavior*, *26*(3), 353–364. https://doi.org/10.1023/A:1015380522722.

Brigham, J. C., & Cairns, D. L. (1988). The effect of mugshot inspections on eyewitness identification accuracy. *Journal of Applied Social Psychology*, *18*(16), 1394–1410. https://doi.org/10.1111/j.1559-1816.1988.tb01214.x.

Cutler, B. L., Penrod, S. D., O'Rourke, T. E., & Martens, T. K. (1986). Unconfounding the effects of contextual cues on eyewitness identification accuracy. *Social Behaviour*, *1*(2), 113–134.

Deffenbacher, K. A., Bornstein, B. H., & Penrod, S. D. (2006). Mugshot exposure effects: Retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law and Human Behavior*, *30*(3), 287–307. https://doi.org/10.1007/s10979-006-9008-1.

Deffenbacher, K. A., & Loftus, E. F. (1982). Do jurors share a common understanding concerning eyewitness behavior? *Law and Human Behavior*, *6*(1), 15–30. https://doi.org/10.1007/BF01049310.

Dysart, J. E., Lindsay, R. C. L., Hammond, R., & Dupuis, P. (2001). Mug shot exposure prior to lineup identification: Interference, transference, and commitment effects. *Journal of Applied Psychology*, *86*(6), 1280–1284. https://doi.org/10.1037/0021-9010.86.6.1280.

Erdelyi, H., & Becker, J. (1974). *Hypermnesia for pictures*, (vol. 171, pp. 159–171).

Egan, D., Pittner, M., & Goldstein, A. G. (1977). Eyewitness identification. Law and Human Behavior, 1(2):199–206.

Garrett, B. (2011). *Convicting the innocent: Where criminal prosecutions go wrong*. Cambridge: Harvard University Press.

Godfrey, R. D., & Clark, S. E. (2010). Repeated eyewitness identification procedures: Memory, decision making, and probative value. *Law and Human Behavior*, *34*(3), 241–258.

Goldstein, J. (2019). *Jailing the wrong man: Mug shot searches persist in New York, despite serious risks*, (p. A1). The New York Times Retrieved from https://www.nytimes.com/2019/01/05/nyregion/nypd-mug-shots-false-identification.html.

Haw, R. M., Dickinson, J. J., & Meissner, C. A. (2007). The phenomenology of carryover effects between show-up and line-up identification. *Memory*, *15*(1), 117–127. https://doi.org/10.1080/09658210601171672.

Hinz, T., & Pezdek, K. (2001). The effect of exposure to multiple lineups on face identification accuracy. *Law and Human Behavior*, *25*(2), 185–198. https://doi.org/10.1023/A:1005697431830.

Hothorn, T., Bretz, F., Westfall, P., & Heiberger, R. M. (2012). Multcomp: simultaneous inference for general linear hypotheses. *R Package Version, 1–2*. https://cran.r-project.org/web/packages/multcomp/index.html.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger, & C. I. M. Fergus (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving memory attributions*. Hillsdale: Lawrence Erlbaum Associates.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. https://doi.org/10.1037/0033-2909.114.1.3.

Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence–accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1304–1316. https://doi.org/10.1037/0278-7393.22.5.1304.

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review*, *119*(1), 80–113. https://doi.org/10.1037/a0025648.

Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, *17*, 349–358.

Memon, A., Hope, L., Bartlett, J., & Bull, R. (2002). Eyewitness recognition errors: The effects of mugshot viewing and choosing in young and old adults. *Memory & Cognition*, *30*(8), 1219–1227. https://doi.org/10.3758/BF03213404.

Mickes, L. (2015). Receiver operating characteristic analysis and confidence–accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory. *Journal of Applied Research in Memory and Cognition*, *4*(2), 93–102. https://doi.org/10.1016/j.jarmac.2015.01.003.

Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied*, *18*(4), 361–376. https://doi.org/10.1037/a0030609.

Nairne, J. S., & Pandeirada, J. N. S. (2008). Forgetting. In Learning and Memory: A Comprehensive Reference (pp. 179–194). Elsevier. https://doi.org/10.1016/b978-012370509-9.00148-0

Palmer, M. A., Brewer, N., Weber, N., & Nagesh, A. (2013). The confidence-accuracy relationship for eyewitness identification decisions: Effects of exposure duration, retention interval, and divided attention. *Journal of Experimental Psychology: Applied*, *19*(1), 55–71. https://doi.org/10.1037/a0031602.

Payne, D. G., & Roediger, H. L. (1987). Hypermnesia occurs in recall but not in recognition. *The American Journal of Psychology*, *100*(2), 145–165. https://doi.org/10.2307/1422400.

Pezdek, K., & Blandon-Gitlin, I. (2005). When is an intervening line-up most likely to affect eyewitness identification accuracy? *Legal and Criminological Psychology*, *10*(2), 247–263. https://doi.org/10.1348/135532505X49846.

Potter, R., & Brewer, N. (1999). Perceptions of witness behaviour-accuracy relationships held by police, lawyers and mock-jurors. *Psychiatry, Psychology and Law*, *6*(1), 97–103. https://doi.org/10.1080/13218719909524952.

R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/.

Roediger, H. L., & Payne, D. G. (1982). Hypermnesia: The role of repeated testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(1), 66–72. https://doi.org/10.1037/0278-7393.8.1.66.

Sauer, J., Brewer, N., Zweck, T., & Weber, N. (2010). The effect of retention interval on the confidence-accuracy relationship for eyewitness identification. *Law and Human Behavior*, *34*(4), 337–347. https://doi.org/10.1007/s10979-009-9192-x.

Schreiber, T. A., & Sergent, S. D. (1998). The role of commitment in producing misinformation effects in eyewitness memory. Psychonomic Bulletin & Review, 5(3), 443-448

Sporer, S. L., Penrod, S., Cutler, B., & Read, D. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*, *118*(3), 315–327. https://doi.org/10.1037/0033-2909.118.3.315.

Steblay, N. K. (2011). What we know now: The Evanston Illinois field lineups. *Law and Human Behavior*, *35*(1), 1–12. https://doi.org/10.1007/s10979-009-9207-7.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Steblay, N. K., & Dysart, J. E. (2016). Repeated eyewitness identification procedures with the same suspect. *Journal of Applied Research in Memory and Cognition*, *5*(3), 284–289. https://doi.org/10.1016/j.jarmac.2016.06.010.

Steblay, N. K., Tix, R. W., & Benson, S. L. (2013). Double exposure: The effects of repeated identification lineups on eyewitness accuracy. *Applied Cognitive Psychology*, *654*(August), 644–654. https://doi.org/10.1002/acp.2944.

Valentine, T., Davis, J. P., Memon, A., & Roberts, A. (2012). Live showups and their influence on a subsequent video line-up. *Applied Cognitive Psychology*, *26*(1), 1–23. https://doi.org/10.1002/acp.1796.

Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology*, *17*(8), 969–993. https://doi.org/10.1002/acp.939.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, *50*(3), 289–335. https://doi.org/10.1016/j.jml.2003.10.003.

Wells, G. L., & Bradfield, A. L. (1998). "Good, you identified the suspect": Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology*, *83*(3), 360–376. https://doi.org/10.1037/0021-9010.83.3.360.

Wells, G. L., Olson, E. A., & Charman, S. D. (2003). Distorted retrospective eyewitness reports as functions of feedback and delay. *Journal of Experimental Psychology: Applied*, *9*(1), 42–52. https://doi.org/10.1037/1076-898X.9.1.42.

Wixted, J. T., Mickes, L., Clark, S. E., Gronlund, S. D., & Roediger, H. L. (2015). Initial eyewitness confidence reliably predicts eyewitness identification accuracy. *American Psychologist*, *70*(6), 515–526. https://doi.org/10.1037/a0039510.

Wixted, J. T., Read, J. D., & Lindsay, D. S. (2016). The effect of retention interval on the eyewitness identification confidence – Accuracy relationship. *Journal of Applied Research in Memory and Cognition*, *5*(2), 192–203. https://doi.org/10.1016/j.jarmac.2016.04.006.

Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. Psychological Science in the Public Interest, 18;(1):10–65.

Yap, M. J., Balota, D. a., Tse, C.-S., & Besner, D. (2008). On the additive effects of stimulus quality and word frequency in lexical decision: Evidence for opposing interactive influences revealed by RT distributional analyses. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(3), 495–513. https://doi.org/10.1037/0278-7393.34.3.495.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- ► Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com

RECEIVED by MSC 4/12/2022 1:07:49 PM

APPENDIX D

Justice Project Eyewitness Identification

RECEIVED by MSC 4/12/2022 1:07:49 PM

THE JUSTICE PROJECT

Eyewitness Identification

A Policy Review

Wrongfully convicted at age 25, Calvin Johnson received a life sentence for the rape of a Georgia woman after four different women identified him. Exonerated in 1999, he walked out of prison a 41-year old man.

The true rapist has never been found.

Introduction 2
Recommendations & Solutions 3
Psychological Factors 4
The Science 5
Benefits & Costs 9
Profiles of Injustice 10
Snapshots of Success 15
Voices of Support 17
Questions & Answers 18
Statistics 19
A Model Policy 20
Literature 24

RECEIVED by MSC 4/12/2022 1:07:49 PM

INTRODUCTION

A measure of fairness and accuracy in the criminal justice system

Eyewitness identification is critical to the apprehension and prosecution of criminals. Eyewitness evidence can also be an important tool for exonerating innocent suspects. Groundbreaking research on eyewitness memory over the past three decades, as well as increasing attention to the problems in the cases of wrongfully convicted individuals, has brought the fallibility of eyewitness memory to the fore.

Eyewitness misidentification is widely recognized as the leading cause of wrongful conviction in the United States, accounting for more wrongful convictions than all other causes combined.[1] Since 1989, DNA evidence has been used to exonerate nearly 200 individuals who were wrongfully convicted. Of those, approximately 75 percent were convicted on evidence that included inaccurate and faulty eyewitness identifications.[2] In some cases, these innocent individuals were misidentified by more than one eyewitness.

In the vast majority of criminal cases, however, DNA or other biological evidence is not available to establish guilt or innocence. Given the persuasive nature of eyewitness evidence, as well as the inherent danger of misidentifications—both in convicting the innocent and allowing the true perpetrator to go free—it becomes imperative that we take stock of the procedures within the control of the criminal justice system that contribute to these problems in order to ensure that the most reliable evidence possible makes it into a courtroom and before a jury.

A number of challenges emerge in pursuit of a more accurate protocol, none more prevalent than an historical lack of communication between scientists and law enforcement.[3] Decades of empirical research have proven that a number of small changes to identification procedures can help improve the accuracy and reliability of eyewitness identifications, and help ensure that the highest quality of eyewitness evidence is collected.

What's more, when put to the test in numerous jurisdictions throughout the country, these reforms have met with real-life success. Thus, it may seem surprising that these reforms have not been implemented in police districts across the board.

While much of the research has been extensively documented and peer-reviewed within the scientific community, and the recommendations for reform are widely accepted by experts in the field, these reforms were initially discussed and developed outside the realm of law enforcement.

Starting in the late nineties, however, leading researchers joined with law enforcement and legal practitioners to bridge the gap and comprehensively address eyewitness identification issues at the intersection of the two fields. As a result, guidelines and best practices for law enforcement were developed with the science in mind.

In October 1999, the Department of Justice released a comprehensive guide for law enforcement on procedures for obtaining more accurate eyewitness evidence.[4] However, there is no current national program or federal agency responsible for educating local departments about these reforms—or in assisting with their practical implementation.[5]

Eyewitness misidentification is widely recognized as the leading cause of wrongful conviction in the U.S., accounting for more wrongful convictions than all other causes combined.

Moreover, as reforms are implemented on a jurisdiction-by-jurisdiction basis in some states, there continues to be little opportunity for sharing information and perhaps even less incentive, given the already overloaded criminal caseloads of police, prosecutors and defenders, and the lack of leadership from the courts or legislature on the issue.

This policy review has been designed to facilitate communication among local law enforcement agencies, policymakers, and others regarding the best practices and methods for enhancing the evidentiary value of correct identifications and at the same time reducing the risk of erroneous identifications. By presenting many of the successful methods employed in local jurisdictions, as well as the science behind them, we hope to create a dialogue around recommendations that will enhance the quality of evidence relied upon in criminal trials, as well as confidence in our system of justice.

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECOMMENDATIONS & SOLUTIONS

Getting it right the first time

A handful of specific improvements have emerged as pragmatic strategies for minimizing eyewitness error. While modernizing identification procedures to incorporate advances in eyewitness memory science requires retooling long-standing lineup methods engrained in police culture, the substantial benefits of implementing the protocol are leading more jurisdictions to update their procedures to catch up with the science.

Because eyewitness evidence, much like trace evidence, is susceptible to contamination, some eyewitness identification procedures actually increase the risk of false identification. By improving these procedures in subtle ways, the actual quality of eyewitness evidence can be improved.

The following recommendations reflect the consensus in the scientific community — confirmed by successful implementation in numerous jurisdictions — as to the procedural changes that will enable law enforcement to extract the most reliable evidence from eyewitnesses for use in a criminal investigation.

These practical changes to the identification process help increase the likelihood of identifying the true culprit while enhancing protections for innocent people accused of crimes.

These reforms are equally effective for photographic lineups and live lineups.

CAUTIONARY INSTRUCTIONS

Prior to presenting the lineup members, the eyewitness should be instructed that the perpetrator may or may not be included in the lineup, and that she should not feel compelled to make an identification.

Cautionary instructions respond, in part, to the tendency of witnesses to make a relative judgment by removing some of the pressure on the eyewitness to choose a suspect when the culprit may not be in the lineup.

EFFECTIVE USE OF FILLERS

Only one suspect should appear in each lineup. In addition, at least five fillers should be included in a photo lineup, and at least four fillers in a live lineup. The fillers should resemble the witness's description of the perpetrator, and the suspect should not unduly stand out.

Fillers, if chosen correctly, allow authorities to judge the reliability of an eyewitness. The effective use of fillers is critical to ensuring that an innocent individual is not identified simply because of the composition of the lineup.

DOCUMENTATION

The identification procedure should be carefully documented. Documentation includes preservation of photos in a photo array or photographs taken of a live lineup, recording all individuals present at the lineup, documentation of the witness's statements regarding the lineup members during the procedure, and, if an identification is made, documentation of the witness's degree of confidence in the identification, in the witness's own words, prior to any feedback from authorities.

Careful documentation of the lineup procedures, including a witness's level of certainty that she has correctly identified the perpetrator, when taken immediately following the identification, helps the jury to assess the eyewitness evidence appropriately and minimizes the effects of reinforcing feedback that can distort the confidence level of an eyewitness between the time of the identification and the trial.

DOUBLE-BLIND ADMINISTRATION

The person who administers the lineup should not know the identity of the suspect. This procedure prevents well-intentioned officials from giving inadvertent clues to the witness as to which person in the lineup is the police suspect.

SEQUENTIAL PRESENTATION

The lineup members should be presented to the witness "sequentially" (one at a time) rather than simultaneously (all at once). Sequential presentation should only occur, however, if the identification procedures comply fully with the double-blind administration recommendation.

Presenting the lineup members one at a time to the witness reduces the tendency for witnesses to engage in "comparison shopping." Rather, an eyewitness must judge whether each lineup member matches her memory of the perpetrator, as opposed to making a relative judgment.

RECEIVED by MSC 4/12/2022 1:07:49 PM

PSYCHOLOGICAL FACTORS

Preventing unreliable evidence in the courtroom

Reliable eyewitness evidence is critical to criminal investigation and prosecution, and it plays a powerful role within the criminal justice system. The repeated discovery of misidentifications contributing to wrongful convictions, however, has prompted inquiries into the nature of eyewitness evidence used to convict criminal suspects, and the problems that arise in utilizing human memory in criminal investigations.

The scientific community has brought the knowledge built through decades of research and experiments to bear on eyewitness identification procedures. Important lessons learned in the laboratory, and in the decades of research devoted to eyewitness memory science, are of enormous value in the legal and law enforcement communities. This substantial body of research has revealed that several natural psychological phenomena can undermine the accuracy of eyewitness identification, and that these psychological factors, left unchecked, can lead to unreliable evidence being presented in the courtroom.

Several natural psychological phenomena can undermine the accuracy of eyewitness identification, and these psychological factors, left unchecked, can lead to unreliable evidence being presented in the courtroom.

LINEUPS AS EXPERIMENTS

Just as trace physical evidence (such as DNA or fingerprints) can be contaminated if it is not collected precisely and carefully, so too eyewitness evidence can be spoiled if it is gathered in ways that do not properly control for known sources of error.[6]

As some researchers have described, a lineup is essentially an experiment designed to test a hypothesis: whether the suspect matches the witness's memory of the perpetrator.[7] Like scientific experiments, careful controls must be put in place to ensure accuracy and prevent the witness's memory from being contaminated or skewed.

Essentially, the lineups as experiments analogy suggests that the logic used to conduct experiments — i.e., isolating variables and implementing careful control conditions — can and should be applied to the legal system when conducting lineups. Using some of the tried and true scientific methods for conducting experiments when conducting a lineup greatly reduces, or in some cases eliminates, the risk of contamination of the data (i.e., eyewitness identification evidence).

RELATIVE JUDGMENT

Relative judgment refers to the natural tendency of a witness to consider lineup participants in comparison with one another, as opposed to a more direct comparison of each lineup member with the witness's memory of the culprit. A witness viewing a lineup will thus tend to identify the person who looks most like the perpetrator in comparison to the other members in the lineup.[8] While, at face value, this process seems unproblematic, it can actually lead to inaccurate and unreliable identifications under certain conditions — namely, when the police suspect is innocent.

The purpose of a lineup is to differentiate innocent suspects from those who actually committed crimes using an eyewitness's memory of an event. Thus, when conducting a lineup, law enforcement officers do not know if a suspect included in a lineup is, in fact, the true perpetrator or simply an innocent person suspected of a crime. If the lineup is full of innocent people (an innocent suspect and a group of innocent fillers), however, relative judgment would mean that an innocent person may be identified, because it is likely that there will always be someone in the lineup who looks more like the person who committed the crime than the other members of the lineup.

Sometimes this person will be a filler, and a witness identification will be dismissed. But sometimes an innocent suspect will be the victim of this tendency toward "comparison shopping," because the witness is always making a relative judgment — the witness is always picking the person who looks closest to the culprit relative to the other lineup members, even if the lineup is full of innocent people.

Take for example a six-person lineup that contains the actual culprit. It has been proven that wit-

RECEIVED by MSC 4/12/2022 1:07:49 PM

nesses who saw the same event will often pick someone out of a lineup when the culprit is removed. In other words, regardless of whether a culprit is in a lineup, witnesses tend to pick the person who looks closest to the culprit, even when the culprit is not present. As leading researchers have noted, "The problem with the relative judgment process, therefore, is that it includes no mechanism for deciding that the culprit is none of the people in the lineup."[9]

MALLEABILITY OF WITNESS CERTAINTY

Traditionally, a witness's self-reported degree of certainty in an identification was considered a good indicator of accuracy. Unfortunately, a great deal of research in recent decades has proven this intuitive assumption false. The level of certainty a witness expresses in her eyewitness testimony does not necessarily correlate with the level of accuracy of the identification. An eyewitness's confidence that she has identified the culprit can fluctuate as a result of factors that occur after the identification and have little to do with memory. This is what is referred to as confidence malleability.[10]

For example, experiments have been conducted in which witnesses were shown a staged crime and asked to identify the culprit from a lineup. The lineup they were shown, however, did not contain the culprit. After the witnesses unknowingly made false identifications, they were then asked their level of confidence. Before doing so, however, some of the witnesses were given various types of reinforcing feedback. Those witnesses who received some confirmation of their false identification, whether the information that a co-witness identified the same individual or some other confirming feedback, were far more confident in their identifications than other witnesses who were given no feedback — despite having given false identifications. These witnesses also distorted and exaggerated certain details, such as how good their view was, how much of an opportunity they had to view the culprit, etc.[11] Our new and better understanding of the influence feedback plays on a witness's self-described level of confidence strongly suggests that measures that control for this influence be adopted in our identification procedures.

THE SCIENCE

Demanding changes in eyewitness identification procedures

Scientific treatments of eyewitness evidence began over 100 years ago, most notably with Harvard Professor Hugo Munsterberg's 1908 book, *On the Witness Stand*.[12] While Munsterberg established that eyewitness evidence was much more fallible than previously thought, his research did not show a way forward. Based on the science of the day, the legal system had no capacity for dealing with these mistakes, and the system could not sort the mistakes from the true identifications.[13] The science, at first, only documented the problem, but it could provide no solutions.

In the late 1970's, however, eyewitness memory scientists began to zero-in on the particular sources of eyewitness error and test revised identification procedures that reduced the risk of mistakes. The guiding principle of this new research was that we must do all we can to ensure good quality evidence on the front end of the process, rather than trying to second guess identifications after the fact. For the research on eyewitness fallibility to be useful, it had to be applied to the criminal justice system in a way that allowed the system to prevent or reduce future mistakes. Scientists thus focused on the ways that the system of collecting eyewitness evidence could *itself* cause mistakes, in hopes that these mistakes could be prevented before they occurred.[14]

The past three decades of eyewitness research and discussion have coalesced around this purpose — preventing false identifications with research-based improvements to the system. Largely, these improvements focus on controlling the suggestiveness of the lineup procedures themselves. A discussion of the science behind these improvements follows.

CAUTIONARY INSTRUCTIONS

Regardless of whether the true perpetrator is in a lineup, an eyewitness may feel pressure to make an identification. Witnesses know that, at the very

RECEIVED by MSC 4/12/2022 1:07:49 PM

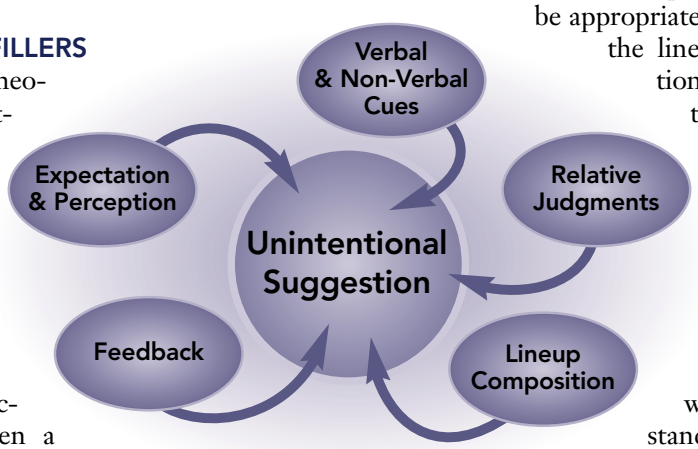
least, a lineup contains a police suspect. When the culprit is not, in fact, present in the lineup, this perception, combined with the natural tendency to compare lineup participants and make a relative judgment, may influence an eyewitness to identify an innocent person.

Cautioning an eyewitness that the offender may or may not be in the lineup reminds witnesses that the answer may be "none of the above."[15]

Research has shown that this extra step, while on its face pointing out a fact that should be obvious, significantly lowers the rate of inaccurate identifications without reducing the number of true identifications.[16]

EFFECTIVE USE OF FILLERS

Relative judgment theory means that an eyewitness viewing a simultaneous lineup tends to make a judgment about which individual in the lineup looks most like the perpetrator relative to the other members of the lineup. This is particularly problematic when a lineup only contains innocent people (i.e., a number of fillers and an innocent suspect).



Research has shown, however, that the effective use of fillers when composing a lineup can help combat the tendency for the relative judgment process to lead to the identification of an innocent suspect.[17]

First, ensuring that the suspect in the lineup does not stand out, or that the fillers resemble the witness's prior description of the culprit at least as much as the suspect does, guards against the eyewitness choosing an innocent suspect simply because the suspect is the only lineup member that resembles the perpetrator.

For example, if the eyewitness describes the perpetrator as an Asian man with a mustache, and there is only one man in the lineup who is Asian and has a mustache, then the lineup is obviously suggestive, and the evidentiary value of any identification is nil. In contrast, if all of the lineup members resemble the prior description of the culprit (or all of the lineup members are Asian men with mustaches), then the eyewitness will have to rely more on comparisons to her own memory of the culprit. In short, no lineup participant can unduly stand out for a lineup to be effective. This holds true in general, but especially with regard to features of the witness's description of the culprit. For example, if a witness describes the perpetrator as having a particular feature such as a mustache, the lineup must be composed with all members sharing that feature.

There are certainly cases where selecting fillers is not as clear-cut. For example, if the suspect does not fit the witness's prior description of the suspect but other evidence creates suspicion of guilt, then it may be appropriate to place that suspect in the lineup, as witness descriptions can sometimes be off the mark. If so, however, the fillers must be chosen to be similar to the appearance of the suspect.[18] There are methods for dealing with contingencies, but the true test of this rule is whether the suspect stands out relative to the other fillers.[19] In other words, if a person who is not involved in the case is given a description of the perpetrator, would she be able to pick the suspect out of the lineup? Including only one suspect in a lineup is also a fundamental safeguard against misidentification. Lineups not only allow police to judge whether a suspect is innocent, they also allow investigators to judge the reliability of an eyewitness. If a lineup contains more than one suspect, however, its ability to test reliability is diminished. This is because it increases the likelihood that a witness would select a suspect based on a guess rather than recognition. The more choices in the lineup test that could be considered "correct" (i.e., suspect identification), the less the lineup can control for witnesses with weak memories or those who guess. The same considerations underpin the need to include an adequate number of fillers. Doing so also reduces the likelihood of an eyewitness identifying an innocent suspect based on a guess. For example, if there is only one suspect and one filler,

RECEIVED by MSC 4/12/2022 1:07:49 PM

the likelihood that an innocent suspect will resemble the culprit more than the other lineup members is 50 percent. If three fillers and one suspect, the likelihood is 25 percent; and so on.[20] While there is no magic number of fillers that should be used, the science has shown that the greater the number of fillers, the greater the reliability of the procedure.

DOCUMENTATION

Lineup identifications are a critical component of the investigation of criminal cases. Given the overwhelming importance of eyewitness testimony and the weight afforded to it by juries, it is essential to provide sufficient contextual information about an identification in order for fact-finders to evaluate its evidentiary weight correctly. Careful documentation of lineup procedures, where possible, means that a complete and accurate record of the methods used to obtain an identification is preserved for review. Recording the identification and the non-identification results, the dialogue between witnesses and police, and the photos themselves (or photographs of a live lineup), serves as much to protect the police from false claims of influencing a witness as it does to preserve the integrity of the evidence. Thorough documentation has the power to put an identification beyond reproach.

A critical component of appropriate documentation is recording an eyewitness's statement of confidence (or self-assessment of certainty) immediately after an identification.

Scientists have shown that a number of procedures within the system actually contribute to misidentifications. Complete documentation allows any suggestiveness in the procedure to be considered by judges and juries in deciding how to weigh the evidence and, when reliable procedures are used, it strengthens the evidentiary value of an identification.

A critical component of appropriate documentation is recording an eyewitness's statement of confidence (or self-assessment of certainty) immediately after an identification. This guards against the confidence malleability problem — when an eyewitness's confidence that she has identified the culprit fluctuates as a result of factors that occur after the identification. To document a witness's confidence, the witness is asked her level of certainty that the person being identified is the true perpetrator *prior to receiving any feedback from authorities or other witnesses*. The witness's confidence level should be recorded in her own words in order to allow judges and juries to evaluate eyewitness testimony in an informed manner.

Studies have shown that information provided to an eyewitness after an identification can influence the witness's level of confidence, and thus skew a juror's assessment of the accuracy of the identification. For example, if an eyewitness makes an identification of a suspect, and that same witness later learns that the person identified also has a criminal record, the witness's confidence level may become artificially inflated.

Confirmatory feedback oftentimes occurs without the knowledge or intent of investigators in the case or even the eyewitness, and if a confidence statement is not taken directly after the identification, the window of opportunity for protecting the integrity of the identification evidence as an indicator of accuracy is lost.

DOUBLE-BLIND ADMINISTRATION

The "double-blind" rule applies the scientific method to lineups, and is rooted in a general strategy for ensuring the objectivity of data collection and interpretation. The purpose of keeping the administrator "blind" as to which person in the lineup is the suspect is to prevent the administrator from unintentionally influencing the results through inadvertent cueing of the witness toward the suspect. A double-blind protocol also eliminates the problem of interpreting ambiguous witness comments and other behaviors through the lens of the theory that the suspect in a lineup is guilty.

Double-blind protocols are familiar to many in the context of pharmaceutical studies to test a new drug. Not only is the patient unaware of whether she received the drug or the placebo, but the doctor who examines the patient during the study is also "blind" to this fact. If the tester knew that the patient had taken the placebo, the tester might unknowingly skew the examination as a result.[21] Double-blind

RECEIVED by MSC 4/12/2022 1:07:49 PM

protocols are standard practice in such contexts not because we distrust the integrity of the medical and scientific professionals involved, but because we understand the risk of natural human psychological factors that can undermine objectivity. We control for such factors because there is much at stake.

Similarly, if a lineup administrator knows which member of the lineup is the suspect, she might unintentionally influence the identification through verbal or non-verbal cues. A cue can be a statement to the witness or even an administrator's posture or facial expression. Verbal and non-verbal cues are examples of suggestive procedures that can suggest to the witness where a suspect is in the lineup.

Verbal and non-verbal cues can also influence or inflate the certainty of the witness. Given that eyewitness confidence is weighed heavily in the legal system, and given that it has been shown to be highly malleable and particularly susceptible to feedback, it is important to design lineup procedures that eliminate the risk of over-inflating confidence through unintentional suggestion.

> The "double-blind" rule applies the scientific method to lineups, and is rooted in a general strategy for ensuring the objectivity of data collection and interpretation.

SEQUENTIAL PRESENTATION

Relative judgment theory also serves as the basis for the sequential presentation recommendation. Traditionally, eyewitnesses are shown a lineup or photo array in which the lineup members are presented as a group. This type of presentation actually encourages an eyewitness to "comparison shop" or to judge the lineup members against each other through a process of elimination.

On the other hand, sequential presentation, first articulated by researchers in the 1980s, is a process where the witness is shown the lineup members one by one and asked to decide if the lineup member presented is the perpetrator. By forcing witnesses to consider the lineup members individually, sequential presentation favors a direct and independent assessment of whether each lineup member matches a witness's actual memory of the perpetrator.

Researchers have shown that the sequential presentation, if implemented in tandem with the double-blind procedure, results in fewer false identifications.[22] It is important to note that if the administrator is not "blind," however, the sequential procedure can actually produce higher rates of false identifications, as witnesses may be more susceptible to unintentional feedback from the administrator when considering one lineup member at a time.

While eyewitnesses have been shown to make fewer choices when viewing a sequential lineup, the research suggests that this is due, in part, to fewer guesses on the part of eyewitnesses with a weak memory of the culprit.[23] A comparison of the accurate and mistaken identifications also suggests that a sequential presentation yields a higher probability that a suspect, if identified using this procedure, is in fact the culprit.[24] In short, the sequential lineup creates a higher threshold for identification by reducing the influence of the tendency to make relative judgments. As a result, the evidentiary value of identifications gained through sequential lineups is much higher, at the cost of some identifications based on weaker witness memory or witness guesses.

Taking this research into the field has shown a generally positive effect. In a pilot project on the sequential procedures in Hennepin County, Minnesota, (Minneapolis) for example, eyewitnesses picked few fillers. Such a low rate of known errors confirmed the value of sequential procedures for officials in that jurisdiction.[25] In addition, New Jersey implemented sequential procedures statewide in 2000. A pilot project conducted by the Chicago Police Department in 2006, however, raised concerns that sequential double-blind lineups were less accurate than conventional methods.[26] Nonetheless, the Chicago study was criticized as having design problems that undermined the study's ability to yield reliable comparisons.[27] Researchers are currently pairing with other jurisdictions to add to the credible literature on the topic. While some questions have been raised about the value of sequential presentation, on balance, most experts believe that it has proven to be superior in both experimental research and in the field.

RECEIVED by MSC 4/12/2022 1:07:49 PM

BENEFITS & COSTS

Investing in a fair and accurate criminal justice system

The benefits of improved eyewitness identification procedures are perhaps best conceived of in terms of the avoided costs. When an eyewitness mistakenly identifies the wrong individual, the costs to public safety are great. Scarce resources in the criminal justice system are misdirected toward investigating and perhaps even trying an innocent person, and not toward convicting and jailing criminals.

"The cost to society of inaccurate eyewitness identifications is twofold," notes psychologist Rod Lindsay of Queens University in Kingston, Ontario. "It's a double error. Not only are you convicting the innocent — or at least putting them through the process of having to get out of the situation — but the guilty are still out there doing the crimes."[28]

A pointed reminder of the costs of misidentification is the case of Clarence Harrison. Wrongfully convicted of a brutal rape in Decatur, Georgia, Clarence Harrison spent nearly 18 years in prison before DNA testing proved his innocence — and showed the eyewitness evidence in his case to be false. After the exoneration, the District Attorney relayed that while the victim was upset by the DNA results, "she is more upset that this means the person who raped her is yet to be identified."

When there are stronger identifications, the benefits to law enforcement and prosecutors, as well as to public safety, are increased, and we can be more confident that the right person is being prosecuted for the crime. In fact, with improved identification procedures, those in law enforcement can ensure that the quality of evidence they are collecting from eyewitnesses is higher than before. Prosecutors are also able to convey to jurors the steps taken to ensure accuracy, placing their evidence on a firm, scientific foundation. And when identifications are better, prosecutions are stronger, and convictions are more solid. By avoiding wrongful convictions, we also avoid the costs of needlessly and unjustly imprisoning an innocent person, as well as the costs of restitution and, in some cases, expensive civil judgments against local governments.

BENEFITS TO INNOCENT SUSPECTS

There can be no question that the conviction of the innocent is a profound injustice. By better protecting the innocent from wrongful conviction, we spare people the devastating ordeal of unjust incarceration that tears apart the families of innocent people and deprives them of their most fundamental liberties. To do justice to our respect for liberty, it is incumbent upon us to do all we can to enhance the accuracy of the criminal justice system.

"LEGAL SAFEGUARDS"

The U.S. Supreme Court has identified five criteria for evaluating the accuracy of eyewitness evidence — "opportunity of witness to view criminal at time of crime, witness' degree of attention, accuracy of witness' prior description of the criminal, level of certainty demonstrated by witness at the confrontation, and length of time between the crime and the confrontation" *(Neil v. Biggers, 1972)*. Unfortunately, although a witness's level of certainty or confidence in her identification is one of the most powerful factors judges and juries consider when assessing eyewitness accuracy, a witness's high level of confidence in an identification does not necessarily mean that the identification is more accurate. In fact, oftentimes the opposite is true.

A number of other procedural protections in place in the legal system to assist in protecting against inaccurate eyewitness evidence have also proven to be starkly inadequate. At last count, more than 75 percent of the nearly 200 wrongfully convicted individuals later exonerated by DNA evidence were convicted on the basis of one or more eyewitness identifications, all with the benefits of legal safeguards to protect against inaccurate identification testimony, such as motions to suppress, cross examination of eyewitnesses and the (limited) admissibility of expert testimony on eyewitness error. Thus, without improvement to the actual quality of the identification procedures themselves, the ability of the legal system to screen out unreliable eyewitness testimony is in grave doubt.

RECEIVED by MSC 4/12/2022 1:07:49 PM

COSTS OF IMPROVED LINEUP PROCEDURES

Reforming eyewitness identification procedures would incur relatively nominal monetary costs or expenditure of departmental time and resources. For example, instructing an eyewitness prior to the lineup, which has been shown to dramatically increase protections for innocent suspects, requires very little training and can be read from a script — it is simply a matter of a change in process. While more careful documentation of identification procedures may seem burdensome, the use of audio or video recording devices can make preservation of the record much easier at nominal cost.

> "The costs of changing procedures are minimal when compared with the benefits. The costs are really a matter of some extra training for our officers. The benefits are stronger, more accurate eyewitness IDs that ultimately make it easier for police and prosecutors to do our jobs."
>
> **John Laux**
> Chief of Police, Bloomington, Minnesota

While some costs may be incurred from implementing a "double-blind" procedure in terms of personnel, alternatives to using an additional officer to administer the lineup can be implemented by using alternative presentation methods that achieve the same result. For example, computer programs that can generate a photo array, and present it to a witness in random order, are increasingly available. Other "low-tech" options include a "folder method," in which the lineup administrator places photos in folders that are shuffled and presented to a witness such that the administrator cannot see the photos while the witness is studying them.[29]

When weighed against the tremendous costs to the taxpayer in terms of lawsuits and compensation to the wrongfully convicted, as well as the very real costs in terms of human lives, the minimal procedural costs associated with these procedures are negligible. Ultimately, the benefits of implementation far outweigh the costs.

PROFILES OF INJUSTICE

Evidence of a broken criminal justice system

Calvin Johnson's Story

Wrongfully convicted at age 25, Calvin Johnson received a life sentence for the rape of a Georgia woman. Four different women identified him. Exonerated in 1999, he walked out of prison a 41-year old man. The true rapist has never been found.

On March 9, 1983, an African-American man entered the apartment of a 30-year old white woman through an unlocked door while she was sleeping in College Park, Georgia. The assailant tightened a belt around her neck until she passed out and then raped her. The victim told police that the attacker turned on the light and that she was able to get a good look at him. Two days earlier, a second woman in College Park had been raped in a remarkably similar manner. College Park straddles the county line between suburban Clayton County and Atlanta's Fulton County. The March 9 attack occurred in Clayton County, and the March 7 rape occurred in Fulton.

THE INVESTIGATION AND EYEWITNESSES

Authorities soon focused on 25-year-old Calvin C. Johnson, Jr., a college graduate recently released from prison for a 1981 burglary. He had readily confessed and pled guilty to the burglary of a College Park man's apartment, and served 20 months. While in jail on that charge, however, police came to suspect Johnson of a sexual assault that occurred the same night of the burglary.

While Johnson was in jail, one of the detectives who worked the burglary went to his cell with a

RECEIVED by MSC 4/12/2022 1:07:49 PM

young white woman. He said they wanted to talk to him about other crimes in the neighborhood, but Johnson refused, telling him he didn't know about any other crimes.

Shortly thereafter, Johnson was charged with the 1981 rape—based on the young woman's identification of *his voice* during the brief jailhouse exchange with the detective. All the rape-related charges were dismissed, however, due to what the prosecutor later characterized as problems with the investigation.

Johnson's lawyer later learned that the victim, who had been forced to have oral sex during the attack, stated several times that her assailant was uncircumcised, a fact that clearly ruled out Johnson. When the College Park rapes occurred two years later, however, suspicions lingered, and Johnson quickly became the target of the investigation.

The same detective from the 1981 cases presented photo arrays, which included Johnson's picture, to both rape victims. The Clayton County victim picked Johnson, but the Fulton County victim picked out another man.

The investigators also showed the photo line-up to two other women who experienced incidents that may have been related to the rapes, as the incidents occurred in the same vicinity and around the same time period. One witness picked Johnson's photo as the man she discovered in her living room when she came out of the shower. The other witness identified his photo as the man who tried to enter her apartment.

The photo of Johnson used in the line-up was from his 1981 arrest, showing him clean-shaven. The perpetrator had been described as clean-shaven, or perhaps having some stubble. At the time of the attacks, however, Johnson had a full beard and moustache — a fact his boss and other witnesses corroborated.

Based on the photo line-ups, police arrested Johnson for rape on March 14, 1983. A search of his home turned up no physical evidence linking him to the crime, but prosecutors later claimed that a jacket of his was similar to one described by one of the victims.

An all-white jury took 45 minutes to find Johnson guilty. On the day of his conviction, he told the judge, "As God is my witness, you've got the wrong man."

Two days after the arrest, detectives arranged a live line-up that included Johnson. This time, with Johnson's lawyer present, the Clayton rape victim did not pick Johnson, but identified a "filler" instead. The two other women who had picked out his photo also failed to pick him out of the live line-up (one identified a filler and the other picked no one). The Fulton rape victim, however, did identify Johnson at the live line-up, though she had failed to identify him from the photo array.

One of the few pieces of physical evidence in the case was a pubic hair found on the Clayton rape victim's sheets. After comparing it with numerous hairs plucked from all over Johnson's body, the state's own forensic experts determined that the hair did not match Johnson. Prosecutors ordered another set of hairs collected from Johnson, but the results were the same—no match.

THE TRIALS

Johnson went to trial for the Clayton County rape on November 2, 1983. Both rape victims identified him in court as their assailant, despite their inconsistent line-up performance.

The two other women who identified Johnson's photo but failed to pick him out of the live line-up also identified him in court as the man from their encounters. As the Fulton rape victim left the witness stand, she lunged at Johnson and cursed him in front of the jury.

Johnson's lawyer presented the testimony of four witnesses who supported his alibi. In addition to the inconsistencies in the photo and live line-up identifications, the defense highlighted the discrepancy between descriptions of a clean-shaven assailant and evidence that proved Johnson had a full beard at the time of the crimes.

The defense also called a state crime lab expert, who testified that the pubic hair found on the victim's bed could not have been Johnson's. The prosecutor argued that the hair must have gotten on the sheet at a public laundry.

After a three-day trial, an all-white jury took 45 minutes to find Johnson guilty. On the day of his

RECEIVED by MSC 4/12/2022 1:07:49 PM

conviction, he told the judge, "As God is my witness, you've got the wrong man." Johnson received a sentence of life plus 15 years.

The following year, Johnson was brought to trial for the rape of the woman in Fulton County. This time, after hearing the same evidence from the same witnesses (plus his conviction for the Clayton County rape), a racially mixed jury unanimously acquitted him. Authorities had virtually no doubt that the same assailant committed both rapes, but the Fulton County acquittal had no effect on Johnson's life sentence.

During those 16 years, Johnson had several opportunities for parole, but the board rejected parole each time because he refused to formally admit guilt and participate in a sex offender program.

THE EXONERATION

With the help of James Bonner, an attorney at the Prisoner Legal Counseling Project at the University of Georgia Law School, Johnson located the evidence from his trial, including a semen sample, though no state law at the time required that the biological evidence be preserved.

According to the prosecutor, when the trial judge retired, a court clerk threw out many old evidence boxes, but someone had pulled Johnson's evidence out of the trash bin and placed it back in storage. In 1994, Johnson wrote to the Innocence Project, and they agreed to take his case.

The Innocence Project arranged to have the remaining evidence sent to Dr. Edward Blake, the nation's foremost forensic DNA expert. Dr. Blake reported that the DNA testing positively excluded Johnson as the source of the semen collected in the rape kit.

Testing of the pubic hair recovered from the victim's sheet also excluded Johnson as the source of the hair, showing a match with the DNA from the rape kit.

On June 15, 1999, the state vacated Johnson's conviction, and Clayton County District Attorney Robert Keller, who had prosecuted the case 16 years earlier, agreed to drop all charges. The true perpetrator has never been found.

In 2000, the Georgia legislature awarded Johnson $500,000 compensation for his wrongful imprisonment.

John Willis' Story

Misidentified by 11 different eyewitnesses for a pattern of crimes involving robbery and rape, John Willis spent over eight years in prison before missing forensic evidence was uncovered that conclusively exonerated him.

Between December 1989 and September 1990, a man the media dubbed the "beauty shop rapist" terrorized the Chatham neighborhood on Chicago's south side.

In the first of a string of remarkably similar and unusual crimes, a man entered a beauty salon brandishing a pistol. He ordered the women in the shop to a back room, forced the women to undress, and robbed them of money and jewelry.

Four crimes of this pattern occurred in beauty shops, and in two of these incidents, on May 2 and September 7, 1990, the man sexually assaulted a female victim. A fifth crime following the pattern of the beauty shop incidents occurred in a store.

THE INVESTIGATION AND EYEWITNESSES

With the help of multiple victim eyewitnesses, police produced and widely distributed a composite sketch.

On September 14, 1990, police arrested John Willis based on an anonymous tip. Though Willis, then 42, had a job cleaning up at a tavern, he had a criminal record of theft and was a self-described "career tire thief and gambler." Willis had no record of violent crime, however, and consistently and emphatically maintained his innocence.

Both of the victims of the sexual assaults identified Willis in photographic lineups as their attacker, as did most of the other witnesses from the salons. A total of 11 eyewitnesses identified Willis as the perpetrator.

THE TRIALS

In 1992, Willis was tried separately for the two crimes that included rapes. In the first case, while no fingerprints or other physical evidence tied Willis to the crime, physical evidence had been col-

RECEIVED by MSC 4/12/2022 1:07:49 PM

lected from the crime scene, including the perpetrator's semen.

The state's forensic analyst from the Chicago Crime Lab, Pamela Fish, testified that her analysis of semen from the crime scene was "inconclusive" — the tests could neither exclude Willis, nor identify him as the source of the semen. In the absence of conclusive forensic testimony, the jury relied entirely on multiple eyewitnesses who had picked Willis out of a lineup, including the rape victim. On February 13, 1992, a jury found Willis guilty, and he was sentenced to 45 years in prison.

While Willis was being held without bond awaiting trial, the string of rapes and robberies continued in Chatham, all following the same unusual pattern.

In April 1992, Chicago Police arrested Dennis McGruder for a string of five rapes and robberies that occurred between November 11, 1991 and March 21, 1992. McGruder pleaded guilty to five crimes that followed the identical pattern of the crimes for which Willis was arrested, including the rape for which he was convicted. One occurred in a beauty salon and four others in taverns.

> Both of the victims of the sexual assaults identified Willis in photographic lineups as their attacker, as did most of the other witnesses from the salons. A total of 11 eyewitnesses identified Willis as the perpetrator.

In November 1993, Willis went to trial for the second rape, after McGruder was jailed for the latter five crimes. A jury again convicted Willis on the basis of identification testimony of the rape victim and other eyewitnesses, along with evidence of the previous rape conviction. Though McGruder had been charged with a string of remarkably similar crimes in the same neighborhood since Willis' arrest, Willis' jury never heard about McGruder.

In an effort to bolster the defense of mistaken identity, Willis' lawyers tried to introduce McGruder's photo into evidence (Willis and McGruder bear a substantial resemblance in their facial features, though Willis is several inches taller than McGruder and noticeably heavier). The prosecution successfully argued to the judge that the McGruder crimes were irrelevant to the case at hand.

At one point during this second trial, Willis became so upset that he tried to blurt out to the jury that the police had someone else in custody for these crimes, waiving a newspaper clipping about the McGruder case in the air. The judge quickly silenced him.

At his sentencing hearing, eyewitnesses from the remaining three crimes with which Willis was charged testified against him. After he was sentenced to an additional 100 years, prosecutors dropped the remaining cases.

In 1997, Cook County Public Defender Greg O'Reilly, the office's leading forensic expert, was brought onto the case to help pursue DNA testing under a new post-conviction DNA statute. When Willis petitioned the court for testing, Assistant State's Attorney Earl Grinbarg, who prosecuted the Willis cases, declared, "John Willis absolutely, positively is the rapist." Nonetheless, Judge Thomas Fitzgerald ordered DNA testing. When O'Reilly sought access to the evidence, he was told that it was all unaccounted for.

An investigation established that Grinbarg had checked the evidence out of the Chicago police evidence room and had not returned the evidence that was not presented as an exhibit at trial, including the biological evidence. The missing evidence—some twenty pieces from three different locations—included swabs taken from the rape victims and a semen-stained toilet paper wrapper, any of which could have been tested for DNA.

Frustrated by the disappearance of the key physical evidence that would allow DNA testing, O'Reilly sought Fish's lab notes. He had been skeptical of her court testimony about inconclusive results and wondered why further testing had not been conducted.

Fish's notes contained evidence of Willis' innocence: they indicated that the blood type of the semen donor of the crime scene evidence was type A, different from Willis' type B. Willis could not have been the source of the crime scene semen.

THE EXONERATION

In September 1998, Willis' lawyers were preparing to appeal based on suppression of the

RECEIVED by MSC 4/12/2022 1:07:49 PM

blood-type exclusion and official misconduct. For 6 months, no biological evidence could be located. Nonetheless, a microscopic slide was eventually discovered in the prosecutor's manila folder among the Willis case files. The slide contained a tiny amount of semen from the first rape for which Willis was convicted.

DNA testing excluded John Willis—and identified Dennis McGruder as the true perpetrator. McGruder was by that time serving a 40 year sentence for the five armed robberies and sexual assaults that occurred after Willis' arrest.

Willis was released on February 24, 1999. He had lost eight and a half years in prison.

At a March 15, 1999 hearing, prosecutors formally dropped all charges against Willis. Thomas Fitzgerald, presiding judge of the criminal division of the Cook County Circuit Court, told Willis, "I wish to God it hadn't happened to you. I hope you can get back on track. And I hope you can live a life that gives you some personal satisfaction and happiness."

The City of Chicago and Cook County settled Willis' civil suit out of court for $2.5 million. He also received $100,000 from the State of Illinois.

Larry Fuller's Story

Larry Fuller spent over 18 years in prison, after being wrongfully convicted of aggravated rape as the result of an erroneous identification — despite the fact that he had a full beard at the time of the identification, which stood in stark contrast to the witness's memory of the perpetrator. Fuller was excluded as the rapist through advanced DNA testing methods, and Governor Rick Perry granted him a full pardon in January 2007.

On the morning of April 26, 1981, a 37-year-old woman was attacked in her apartment in Dallas, Texas by a black man wielding a knife. He cut her several times, raped her, and then ran away. The victim was taken to the hospital, and a rape kit was collected. The attack occurred 45 minutes before sunrise, and the victim testified that it was dark in the room, although she was able to ascertain that the attacker was a black male "somewhere in his twenties" and that she had never seen him before. She also reported that she did not remember any facial hair on the attacker.

THE INVESTIGATION AND EYEWITNESS

At the time of the April rape, Fuller was a decorated Vietnam War veteran raising two young children with his girlfriend while pursuing an education. While he had served three years for robbery after his return from Vietnam, Fuller had no record of sex crimes.

Nonetheless, investigators had obtained Fuller's photograph while investigating an incident that occurred three months earlier. In this previous incident, on the morning of January 19, 1981, another woman had been similarly raped, just a few buildings down from the victim of the April rape.

Fuller was stopped by police after the earlier attack because he matched the victim's description, but when his photograph was placed in a photo array, the victim positively stated that the photo array did not contain her attacker. Another man, Larry James Johnson, later confessed to the January crime and was arrested and prosecuted.

Two days after the April attack, police included Larry Fuller's picture in a photo array presented to the victim of the April attack. The victim failed to conclusively identify Fuller as her attacker, however, telling investigators Fuller "looks a lot like the guy" but she could not identify him. The investigating officer then issued a report recommending that the investigation be suspended because the victim "was unsure of the suspect at this time."

Five days after the first photo lineup, on May 3, 1981, police showed the victim a second photo array, this time with a more recent picture of Fuller, taken the same day. *Fuller's was the only photograph included in both photo arrays.*

This time, the victim positively identified Fuller, though she was confused by the fact that Fuller had a heavy and distinct beard. She had stated previously that she did not remember any facial hair on the attacker.

The victim later stated, "I looked at it, and I knew that was the face; but I couldn't figure out why there was facial hair because I didn't remember the facial hair . . . I looked at the picture again and I put my finger over the part, the hair, and then I could identify him."

RECEIVED by MSC 4/12/2022 1:07:49 PM

THE TRIAL

At trial, the prosecution relied on the eyewitness identification, stating that the victim "never wavered" in her identification, and the victim testified that she was certain Fuller was her attacker.

In addition, the prosecution introduced complex expert testimony on serological testing of semen from the rape kit collected from the victim. Though technology at the time did not allow for advanced DNA testing, the Southwestern Institute of Forensic Sciences did perform more basic tests on the semen evidence. A forensic serologist testified that Fuller could have been the source of the semen based on this testing, but it was inconclusive. The prosecution incorrectly argued that the semen evidence was consistent with Fuller's to the exclusion of 80 percent of the population — a major exaggeration of the evidentiary value of the testing.

The defense called Fuller's girlfriend, who testified Fuller was at their house at the time of the attack. Despite his alibi, on August 25, 1981, Fuller was convicted of aggravated rape after only 35 minutes of jury deliberation. He received a sentence of 50 years in prison on September 10, 1981.

THE EXONERATION

Fuller wrote the Innocence Project in the mid 1990s, and they agreed to help him pursue more advanced DNA testing of the physical evidence. Meanwhile, in 1999, after having served 18 years of his sentence, Fuller was released on parole, but was sent back to prison in 2005 for a parole violation.

In November 2000, the Innocence Project located the biological evidence at Southwestern Institute of Forensic Sciences and requested that the Dallas County District Attorney's Office consent to post-conviction DNA testing. In March 2001, the Office refused, noting that the Texas legislature was considering a new DNA statute, and they wanted to wait for the statutory criteria.

In August 2001, the Innocence Project again requested testing under Texas' new post-conviction DNA statute, but the state opposed testing. However, after a hearing in judicial court, the judge ordered that DNA testing be conducted by the Department of Public Safety (DPS).

Unfortunately, DPS was unable to obtain the profile of the male DNA on the vaginal slide, and in November 2004, the Innocence Project renewed its request to the District Attorney's Office for more-developed DNA testing using another method. On April 14, 2006, the District Attorney's Office agreed, and the Court ordered Y-STR testing at Orchid Cellmark, a private laboratory.

Having waited a quarter of a century, Fuller received unassailable proof of his innocence — Y-STR testing conclusively excluded him as the source of the semen.

At a hearing on October 31, 2006 in the 203rd Judicial District Court in Dallas, Judge Lana McDaniel released Fuller. Although not involved in the original case, the judge said she felt sick to her stomach over the time he spent in prison for a crime he did not commit.

On January 25, 2007, Fuller received a full pardon from Texas Governor Rick Perry. He was the tenth person from Dallas County to be exonerated by DNA evidence in the last five years.

SNAPSHOTS OF SUCCESS

If it works in these states and jurisdictions, why not the rest of the country?

NEW JERSEY

The first state in the nation to officially adopt the National Institute of Justice recommendations issued in 1999 *(Eyewitness Evidence: A Guide for Law Enforcement)*, New Jersey provides an example of the successful implementation of reform protocols and their pragmatic effectiveness. While most law enforcement agencies or departments are controlled locally, the Attorney General of New Jersey was able to mandate changes in procedure across the entire state due to the unique supervisory authority of the Attorney General in that state. Since April 2001, New Jersey has conducted double-blind, sequential lineups. In addition, police officers issue cautionary instructions, ensure that lineups are constructed effectively with an adequate number of appropriate

RECEIVED by MSC 4/12/2022 1:07:49 PM

fillers, and document the identification procedures, including the witness's statement of certainty.

On July 31, 2006, the New Jersey Supreme Court, noting the importance of a complete record of an identification procedure in ensuring the reliability of eyewitness evidence presented to a jury, made complete documentation of the identification procedure a condition of admissibility of out-of-court identifications.[30] According to the opinion, "[G]iven the importance of ensuring the accuracy and integrity of out-of-court identifications, we will exercise our rulemaking authority to require . . . that the police record, to the extent feasible, the dialogue between witnesses and police during an identification procedure." The decision was unanimous.

NORTH CAROLINA

In November 2002, Justice I. Beverly Lake created the North Carolina Actual Innocence Commission to study and recommend potential strategies for lessening the incidence of wrongful convictions. The Commission issued recommendations for eyewitness identification in October 2003 and endorsed changes in procedures such as the delivery of cautionary instructions, documentation of a witness's confidence in the identification without any feedback given by the administrator, effective use of fillers (a minimum of eight photos in photo identification procedures and a minimum of six individuals in live identification procedures), and sequential double-blind presentation.[31]

While the Commission has no official authority over law enforcement agencies in the state, the Commission members include the North Carolina attorney general, district attorneys, police chiefs, Supreme Court Justices, and others. A number of North Carolina's law enforcement agencies are increasingly implementing the Commission's eyewitness recommendations to date.

WISCONSIN

After studying the problem of mistaken identifications, the Training and Standards Bureau of the Wisconsin Department of Justice, working with the University of Wisconsin Law School's Frank J. Remington Center, developed a comprehensive set of eyewitness identification guidelines for law enforcement, which were adopted and distributed to law enforcement throughout the state in March 2005.[32] The guidelines — which include cautionary instructions to eyewitnesses, assessments of confidence immediately after identifications, proper selection of fillers, and double-blind, sequential presentation of lineups — represent a model for implementation of the "best practices" in eyewitness identification.

Legislation passed in November 2005 requires each law enforcement agency in the state to adopt policies or guidelines on eyewitness identification procedures. Though the model policy developed by the Attorney General is not mandatory, the Wisconsin Department of Justice has developed a training program to educate law enforcement across the state on the need for changes in procedure to lessen the risk of misidentification. Some departments have adopted the model policy, and more are likely to follow. To date, the program has trained over 800 investigators on the new procedures. Training on these procedures has also been incorporated into the curriculum for new investigators.[33]

MINNESOTA

Beginning in 2003, Hennepin County Attorney Amy Klobuchar spearheaded an effort to implement a sequential, double-blind pilot program in four police departments in the state, including Minneapolis. A follow-up study analyzing the pilot found that the pilot project was relatively easy to implement, with projects up and running in the smaller counties in two weeks, and in the larger counties in under a month. The reforms incurred minimal costs, no perceived drop in suspect identifications, and a reduction in filler identifications.[34] The study showed increased protections against misidentification, practical benefits for investigators, and a higher quality of eyewitness evidence. As a result of the pilot, the Hennepin County Attorney urged adoption of the reform protocol county-wide.

OTHER STATES

In 2003, the Illinois legislature passed legislation mandating cautionary instructions, as well as documentation and lineup composition requirements. In addition, a number of individual jurisdictions throughout the country have adopted reforms at the local level. These jurisdictions include the Boston Police Department and other departments in Suffolk County (in coordination with the Suffolk County District Attorney), Northampton, Massachusetts, Virginia Beach, Virginia and Santa Clara, California, among others.[35]

RECEIVED by MSC 4/12/2022 1:07:49 PM

VOICES OF SUPPORT

"I did see many flaws in witnesses who felt like they were trying to be people pleasers, felt they had to select someone. Now people are actually comparing the one photo in front of them to what's in their mind, not going through process of elimination."[36]

Joy Rikala
Chief of Police
Minnetonka Police Department
Governing Magazine, May 2006

"It is axiomatic that eyewitness identification evidence is often crucial in identifying perpetrators and exonerating the innocent. However, recent cases, in which DNA evidence has been utilized to exonerate individuals convicted almost exclusively on the basis of eyewitness identifications, demonstrate that this evidence is not fool-proof . . . While it is clear that current eyewitness identification procedures fully comport with federal and state constitutional requirements, the adoption of these Guidelines will enhance the accuracy and reliability of eyewitness identifications and will strengthen prosecutions in cases that rely heavily, or solely, on eyewitness evidence."[37]

John J. Farmer, Jr.
New Jersey Attorney General
Memorandum, April 18, 2001

"Every time you see something coming along that makes your job a little harder, you kind of cringe a little. It's going to take extra time and personnel, but if it's going to make a case a little more solid or if it's going to eliminate a bad identification or a situation where an officer may try to influence an identification, then it's beneficial."[38]

John E. Miliano
Chief of Police, Linden, New Jersey
New York Times, July 21, 2001

"If you don't do this, you risk having good convictions and good identifications thrown out."[39]

David Angel
Deputy District Attorney
Santa Clara County, California
Pittsburgh Post-Gazette, May 9, 2005

"We hadn't changed the way we do eyewitness procedures in decades . . . DNA [exonerations] obviously have shown us that we have to change."[40]

Ken Patenaude
Detective Lieutenant
Northampton, MA Police Department
Pittsburgh Post-Gazette, May 9, 2005

"The psychology behind these procedures is to have witnesses focus on their actual memory of the incident and the suspect. We want to eliminate any kind of extraneous influence or bias in the identification process."[41]

Robert Olson
Chief of Police, Minneapolis, Minnesota
November 3, 2003

"I will never forget the day I learned about the DNA results. I was standing in my kitchen when the detective and the district attorney visited. They were good and decent people who were trying to do their jobs — as I had done mine, as anyone would try to do the right thing. They told me: "Ronald Cotton didn't rape you. It was Bobby Poole." The man I was so sure I had never seen in my life was the man who was inches from my throat, who raped me, who hurt me, who took my spirit away, who robbed me of my soul. And the man I had identified so emphatically on so many occasions was absolutely innocent . . . If anything good can come out of what Ronald Cotton suffered because of my limitations as a human being, let it be an awareness of the fact that eyewitnesses can and do make mistakes."[42]

Jennifer Thompson
Victim/Activist for Eyewitness Identification Reform
New York Times, June 18, 2000

"God forbid that we would put an innocent person in jail because of a less than confident eyewitness. And then we would be allowing a guilty person to go out and commit more crimes."[43]

William Mullen
Chief Deputy Sheriff of Allegheny County and
Former Assistant Chief of Police, Pittsburgh, PA
Associated Press, May 9, 2005

RECEIVED by MSC 4/12/2022 1:07:49 PM

QUESTIONS & ANSWERS

Why change our existing protocol, which has worked for years?

Given the firm scientific basis for recommending the protocol, it is worth comparing these justifications with the current standard protocol. The standard way of conducting lineups today is not rooted in careful science. Rather, it was developed as an *ad hoc* procedure created and embraced in the law enforcement community because of its intuitive plausibility. Nothing more recommends or justifies it than tradition.

Nonetheless, eyewitness memory science has established that many factors related to eyewitness memory that seem intuitive and obvious are not necessarily true. For example, a witness's confidence in an identification is not a reliable predictor of accuracy. This is counter-intuitive, but the lack of a close correlation has been very well documented. Traditional methods need to be updated and procedures modernized to catch up with our modern understanding of eyewitness memory issues.

While no one can deny that many guilty people have been convicted based on evidence obtained with the traditional protocol, we have witnessed far too many innocent people convicted based on incorrect eyewitness testimony and later exonerated by DNA testing. The result — investigations led off course and prematurely ended, allowing predators to go uninvestigated and unpunished. It is incumbent upon us to live up to our commitment to public safety and base our procedures on the best science, not tradition.

Why haven't we heard of the research or improved procedures before?

It is not surprising that many people in law enforcement are unfamiliar with this scientific research. Police officers typically do not read technical peer reviewed academic journals (who could blame them?) or attend conferences about experimental psychology. Increasingly, however, opportunities have been created to foster a dialog, and many law enforcement agencies have modernized procedures based on the science. Only in recent years, upon the dawning of the age of DNA, have people begun to appreciate the problem of mistaken eyewitness identification, leading people in all aspects of the criminal justice system to look more carefully at ways of enhancing accuracy and putting higher quality evidence into the courtroom.

Isn't the blind administration component an insult to the integrity and professionalism of detectives?

Requiring a neutral administrator is NOT about challenging the integrity or professionalism of law enforcement personnel. Structuring procedures to generate the best quality of evidence is what professionalism demands. The issues addressed here have nothing to do with suggestions of misconduct. Rather, they address certain realities about normal human psychology and the possibility of the *inadvertent* cuing of a witness. All manner of verbal and nonverbal human behaviors may have the *unintended* effect of influencing a witness. Using a neutral administrator eliminates this possibility and ensures the best quality of evidence.

Just as in double-blind clinical drug trials, we are not assuming doctors and medical researchers are nefarious and dishonest; requiring neutral administrators is simply good practice — especially with such important matters as liberty and public safety on the line.

The courts don't seem to have any problems with the standard procedures, so why change?

The courts have increasingly begun to recognize that many of our traditional assumptions about eyewitness memory are wrong (such as the link between certainty and accuracy). Exonerations have made clear the need for change (and the terrible human costs of persisting with traditional practices), and developments in eyewitness memory science have identified ways of enhancing accuracy through more carefully designed procedures. Because the state of the science is now very solid, courts have often been more willing to allow challenges to existing protocols. Rather than picking apart in-court identifications that follow from flawed procedures, it is in the best interest of all parties to implement best practices that guarantee the best quality of evidence at the outset, on the front end of the process. Some courts have already ordered new procedures on that basis.

Why should we care about experiments with "staged crimes" and "mock witnesses"?

Experimental psychologists carefully design their experiments to isolate certain phenomena so that they may be better observed and understood in ways that

RECEIVED by MSC 4/12/2022 1:07:49 PM

'real world' observation does not allow. These methods are the only way to fully control the different variables and track their changes under different conditions. In actual cases, for example, we cannot otherwise be completely certain whether an identified suspect is, in fact, the perpetrator the way we can in experiments.

While the experiments have created a solid basis for the various elements of the protocol, we know from real world applications (statewide in New Jersey, as well as in many other individual jurisdictions) that the system is practical and pragmatically workable. There is already a track record of real world success.

Is the protocol practically feasible, especially for some smaller departments?

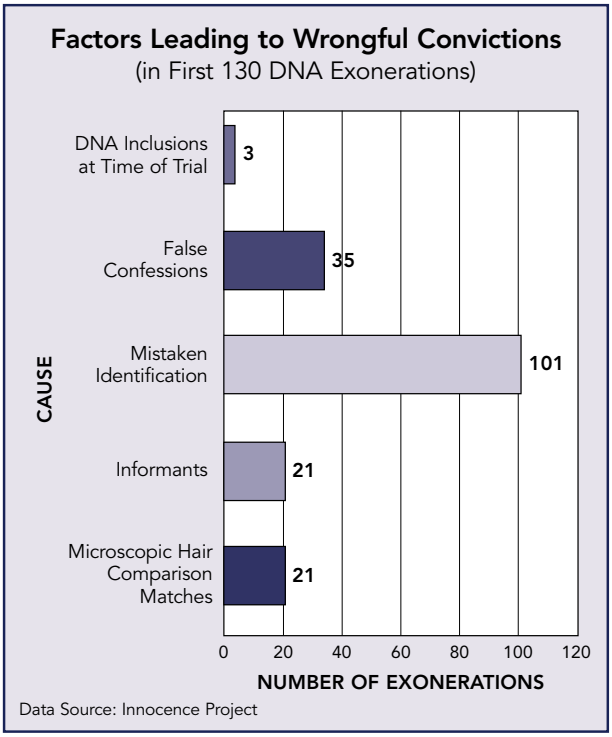
Experience in other jurisdictions across the country shows that the protocol is practical and workable. The protocol is sensitive to the potential problem of finding a neutral administrator, and provides for alternatives that accomplish the same goals.

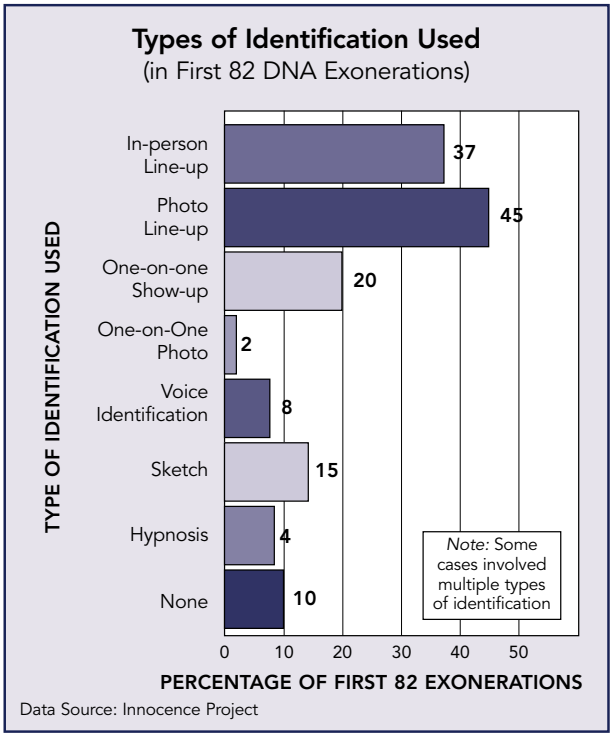
STATISTICS

The Innocence Project found that mistaken eyewitness identifications were the leading cause of the first 130 DNA exonerations, accounting for 101 of the total. A subsequent study by the Innocence Project found that over 75 percent of the now nearly 200 post-conviction DNA exonerations in the U.S. involve mistaken eyewitness identification testimony, making it the leading cause of these wrongful convictions.

In addition, the Center on Wrongful Convictions at Northwestern University School of Law also studied the cases of 86 defendants who had been sentenced to death but legally exonerated based on strong claims of actual innocence, finding that eyewitness testimony played a role in the convictions of 54 percent of the death-sentenced defendants. Eyewitness testimony was the only evidence used against 38 percent of the defendants.

The Innocence Project also found that photo lineups were the most oft-used identification method in the first 82 DNA exonerations. Investigators used a photo lineup in 45 percent of the cases.





RECEIVED by MSC 4/12/2022 1:07:49 PM

A MODEL POLICY

AN ACT CONCERNING EYEWITNESS IDENTIFICATION PROCEDURES

WHEREAS the goal of a police investigation is to accurately identify and apprehend the true perpetrators of crimes; and

WHEREAS eyewitness error is the leading cause of mistaken convictions; and

WHEREAS cases of mistaken conviction in [insert jurisdiction] owing to eyewitness misidentification have resulted in the actual perpetrators remaining free to commit more crimes; and

WHEREAS scientific studies of eyewitness memory have demonstrated that eyewitness evidence is, like trace physical evidence, susceptible to contamination if not handled properly; and

WHEREAS well-intentioned witnesses and authorities acting in good faith may sometimes inadvertently undermine the accuracy of an identification procedure unless appropriate safeguards are in place; and

WHEREAS extensive scientific research has shown that modified methods of conducting identification procedures greatly enhance accuracy;

We hereby enact the following

EYEWITNESS ACCURACY ENHANCEMENT ACT:

Section 1: Definitions. For purposes of this section the following definitions apply:

1) Photo Lineup: a selected group of photographs of persons presented to an eyewitness to a crime, containing a single suspect and several fillers, for the purpose of determining whether the eyewitness is able to identify the suspect as the perpetrator.
2) Live Lineup: A selected group of persons presented to an eyewitness to a crime containing a suspect and several fillers for the purpose of determining whether the eyewitness is able to identify the suspect as the perpetrator.
3) Suspect: A person under investigation for participation in a crime.
4) Filler: A person, not a suspect in the crime under investigation, not known to the witness, who is made part of a live lineup; or a photograph of a person, not a suspect in the crime under investigation, not known to the witness, made part of a photo lineup and presented to a witness.
5) Neutral Blind Administrator: A person who conducts photo or live lineup procedures while unaware of which person in the lineup is the suspect and which are fillers.

Section 2: Development and Dissemination of Eyewitness Identification Protocol. Prior to [insert date] the [insert jurisdiction] Attorney General shall consult with law enforcement and scientific experts in eyewitness memory to develop, adopt, and disseminate to all law enforcement jurisdictions in the state comprehensive, written policies and procedures and associated training materials for [insert jurisdiction] law enforcement agencies regarding photo and live lineup eyewitness identification procedures that implement the requirements set forth in section 3 of this act.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Section 3: Requirements for Photo and Live Line-up Identification Procedures.
For any offense alleged to have been committed on or after [insert date], all photo and live lineup identification procedures conducted by law enforcement officers shall be administered pursuant to the procedures developed by the Attorney General described in section 2 of this act and consistent with the requirements in this section.

A. Witness Instructions. Prior to presenting a live lineup or photo array identification procedure, the lineup administrator shall instruct the witness that:

1) The procedure is intended to identify guilty parties as well as to clear innocent suspects from suspicion;
2) The witness should not guess or conclude that the perpetrator is among the persons in the lineup;
3) The witness should not feel compelled to make an identification because the perpetrator may or may not be among those shown;
4) The person administering the lineup may not be aware of which person in the lineup is the suspect;
5) Individuals depicted in lineup photos may not appear exactly as they did on the date of the incident because features such as head and facial hair are subject to change;
6) The police will continue to investigate the incident whether or not the witness identifies someone.

B. Documentation of Identification Procedures.

1) All photo and live lineup identification procedures conducted in connection with a criminal investigation shall be documented, regardless of whether an identification is made, made a part of the case record, and provided to the prosecuting authority in the event any prosecution related to the crime being investigated occurs. The documentation shall include:
    a. The time, date, location and identities of all persons present;
    b. A form listing the instructions enumerated in section A. of this act signed by the witness to confirm understanding of the instructions prior to administration of the identification procedure;
    c. A photograph of any live lineup as presented to a witness; or all photographs used in any photo lineup preserved in their original condition;
    d. The order of presentation of photographs or individuals.
2) All comments and exchanges during an identification procedure shall be electronically recorded with audio or audio/video recording equipment whenever possible. When it is not feasible to electronically record the identification procedure, comments and exchanges among persons present during an identification procedure shall be documented in writing, and an explanation of why electronic recording was not feasible shall be included in the record. The documentation, whether electronic or written, shall include all witness comments, using the witness's own words, regarding the persons or photos in the lineup and all questions and commentary by the lineup administrator and any other persons present during the identification procedure.
3) If the witness makes an identification as a result of a photo or live lineup, the lineup administrator shall immediately ask the witness to state in his or her own words how confident he or she is that

RECEIVED by MSC 4/12/2022 1:07:49 PM

the person identified is the perpetrator, and make the witness's words part of the record prior to any commentary or feedback from the lineup administrator or any other persons present.

4) If no electronic recording of the identification procedure is made, the witness shall review and sign the written record of the identification procedure, including all comments regarding the persons or photos presented, and any statements regarding an identification and degree of certainty, prior to any feedback or communication of information from the administrator or others involved in the investigation regarding the identification procedure.

C. General requirements for composition and conduct of lineup identification procedures

1) During the identification procedure, the administrator shall refrain from any commentary or feedback to the witness regarding particular persons or photographs in a lineup until after the procedure is concluded and the witness certifies the record of the procedure.
2) At least five fillers shall be included in a photo lineup, in addition to the suspect, and at least four fillers shall be included in a live lineup, in addition to the suspect.
3) Only one member of a photo or live lineup shall be a suspect, and the remainder shall be fillers who are not suspects.
4) Fillers shall be selected who generally fit the witness's description of the perpetrator. When there is a limited or inadequate description of the perpetrator provided by the witness, or when the description of the perpetrator differs significantly from the appearance of the suspect, fillers should resemble the suspect in significant features.
5) Lineup administrators shall create a consistent appearance between the suspect and fillers with respect to any unique or unusual feature such as scars or tattoos used to describe the perpetrator by artificially adding or concealing that feature in filler photographs.
6) In photo line-ups, the suspect's photo should resemble his or her appearance at the time of the offense and not unduly stand out.
7) If the eyewitness has previously viewed a photo lineup or live lineup in connection with the investigation of another person suspected of involvement in the offense, the fillers in the lineup in which the suspected perpetrator participates shall be different from the fillers used in any prior lineups.
8) Law enforcement shall seek identification of any particular suspect through photo or live lineup only once from any given witness.
9) In a photo lineup, no writings or information concerning any previous arrest, indictment, or conviction of the suspected perpetrator shall be visible or made known to the eyewitness.
10) The position of the suspect in a photo or live lineup should be changed for each new witness to view the photo lineup.
11) In a live lineup, any identifying actions, such as speech, gestures, or other movements, shall be performed by all lineup participants.
12) In a live lineup, witnesses shall not be exposed to the members of the lineup before the procedure begins.

D. Neutral Blind Administration of Photo and Live Line-ups

1) Whenever possible, the administrator of photo or live lineup identification procedures shall be someone who is not aware of which member of the lineup is the suspect in the case and which

RECEIVED by MSC 4/12/2022 1:07:49 PM

are fillers, and no person familiar with the identity of the suspect shall be present during the identification procedure.

2) When it is not feasible to have the procedure administered by someone unaware of which person is the suspect, that reason shall be documented, and a photo lineup procedure may be conducted using an alternative method specified and approved by the Attorney General. Any alternative procedure shall be structured to achieve neutral blind administration and prevent the administrator from viewing the lineup simultaneously with the witness or knowing the order of photographs as presented to the witness during the identification procedure. Alternative methods may include the following:
    i. automated computer programs approved by the Attorney General for such use that can automatically administer the lineup identification procedure directly to a witness, and during which the administrator cannot see which photo the witness is viewing until after the procedure is completed; or, alternatively,
    ii. a procedure approved by the Attorney General in which photographs are placed in folders, randomly numbered and shuffled, and then presented to a witness such that the administrator cannot see or determine the order of photograph being presented to the witness until after the procedure is completed; or, alternatively,
    iii. other such procedures as specified by the Attorney General which achieve neutral blind administration.

Note: Due to a lack of comprehensive data from pilot studies, the above model does not include a provision regarding sequential procedure. Nonetheless, researchers are currently pairing with other jurisdictions to add to the credible literature on the topic. While some questions have been raised about the value of sequential presentation, on balance most experts believe that it has proven to be superior in both experimental research and in the field. Thus, jurisdictions may also want to consider the addition of the sequential procedure, if and only if, neutral-blind administration is employed. In that event, the following provision may be inserted in the above model:

E. Sequential Procedure.

1) Live line-up and photo array identification procedures shall be presented to witnesses using a sequential method, in which a witness is shown photographs or live lineup participants one at a time, and not simultaneously. The witness shall be asked to state for each person whether the individual shown is the perpetrator, prior to viewing the next lineup participant.
2) The administrator shall not offer any comment or feedback to the witness regarding the witness's responses.
3) If there are multiple eyewitnesses, witnesses shall be presented with the identification procedure separately, and the suspect shall be placed in a different position in the lineup for each eyewitness.
4) Under no circumstances shall a sequential presentation be used unless the procedure complies fully with neutral blind administration specified in section D.

RECEIVED by MSC 4/12/2022 1:07:49 PM

LITERATURE

SUGGESTED READINGS

The following materials are essential reading for individuals interested in improving eyewitness identification procedures.

Klobuchar, A., N.K.M. Steblay, and H.L. Caligiuri. "Improving Eyewitness Identifications: Hennepin County's Blind Sequential Lineup Pilot Project." Cardozo Public Law, Policy & Ethics Journal 4 (2006): 381-413. http://web.augsburg.edu/~steblay/Improving_Eyewitness_Identifications.pdf

Wells, G.L., M. Small, S.D. Penrod, R.S. Malpass, S.M. Fulero, and C.A.E. Brimacombe. "Eyewitness Identification Procedures: Recommendations for Lineups and Photospreads," *Law and Human Behavior* 22 (1998): 603-647. http://www.psychology.iastate.edu/FACULTY/gwells/Wells_articles_pdf/whitepaperpdf.pdf

Wisconsin Office of the Attorney General. Training and Standards Bureau. *Model Policy and Procedures for Eyewitness Identification* (2005). http://www.doj.state.wi.us/dles/tns/eyewitness.asp

SELECTED BIBLIOGRAPHY

The following listing includes some of the key source material used in developing the content of this policy review. While by no means an exhaustive list of the sources consulted, it is intended as a convenience for those wishing to engage in further study of the topic of improved eyewitness identification procedures. Many of the entries contain hyperlinks for ease in locating an article, report or document on the web.

1. Journals and Law Reviews

Bradfield, A.L., G.L. Wells, and E.A. Olson. "The Damaging Effect of Confirming Feedback on the Relation between Eyewitness Certainty and Identification Accuracy." *Journal of Applied Psychology* 87 (2002): 112-120.

Gross, S.R., K. Jacoby, D.J. Matheson, N. Montgomery, and S. Patil. "Exonerations in the United States, 1989 Through 2003." *The Journal of Criminal Law & Criminology* 95 (2005): 523- 560. http://www.thejusticeproject.org/press/reports/pdfs/17220.pdf

Malpass, R.S. "A Policy Evaluation of Simultaneous and Sequential Lineups." *Psychology, Public Policy and Law* 12, no. 4 (2006): 394-418. http://eyewitness.utep.edu/Documents/Malpass06PolicyEvaluationOfSimultaneousAndSequentialLineups-PPPL.pdf

Penrod, S. D., and B.L. Cutler. "Witness confidence and witness accuracy: Accessing their forensic relation." *Public Policy, Psychology & Law* 1 (1995): 817-845.

Semmler, C., N. Brewer, and G.L. Wells. "Effects of Postidentification Feedback on Eyewitness Identification and Nonidentification." *Journal of Applied Psychology* 89, no. 2 (2004): 334-346. http://www.psychology.iastate.edu/faculty/gwells/Wells_articles_pdf/JEPA_Semmler_et_al.pdf

Steblay, N.M. "Social Influence in Eyewitness Recall: A Meta-Analytic Review of Lineup Instruction Effects." *Law and Human Behavior* 21, no. 3 (1997): 283-297. http://web.augsburg.edu/%7Esteblay/LineupInstructionAnalysis.pdf

RECEIVED by MSC 4/12/2022 1:07:49 PM

Steblay, N., J. Dysart, S. Fulero, and R.C.L. Lindsay. "Eyewitness Accuracy Rates in Sequential and Simultaneous Lineup Presentations: A Meta-Analytic Comparison." *Law and Human Behavior* 25, no. 5 (2001): 459-473. http://web.augsburg.edu/%7Esteblay/EyewitnessAccuracy.pdf

Turtle, J., R.C.L. Lindsay, and G.L. Wells. "Best Practice Recommendations for Eyewitness Evidence Procedures: New Ideas for the Oldest Way to Solve a Case." *The Canadian Journal of Police & Security Services* 1 (2003): 5-18. http://www.ryerson.ca/~jturtle/cjpss.html

Wells, G.L. "Applied Eyewitness-Testimony Research: System Variables and Estimator Variables." *Journal of Personality and Social Psychology* 26, no. 12 (1978): 1546-1557.

Wells, G.L. Eyewitness Identification: Systemic Reforms, *Wisconsin Law Review* 2006, no. 2 (2006): 615-643. http://www.psychology.iastate.edu/FACULTY/gwells/Wisconsin_Law_article.pdf

Wells, G.L., and A.L. Bradfield. "'Good, You Identified the Suspect': Feedback to Eyewitnesses Distorts their Reports of the Witnessing Experience." *Journal of Applied Psychology* 83 (1998): 360-376.

Wells, G.L., M. Small, S.D. Penrod, R.S. Malpass, S.M. Fulero, and C.A.E. Brimacombe. "Eyewitness Identification Procedures: Recommendations for Lineups and Photospreads," *Law and Human Behavior* 22, no. 6 (1998): 603-647. http://www.psychology.iastate.edu/FACULTY/gwells/Wells_articles_pdf/whitepaperpdf.pdf

2. Commission and Association Reports & Policies

American Bar Association. "Achieving Justice: Freeing the Innocent, Convicting the Guilty." *Report of the ABA Criminal Justice Section's Ad Hoc Innocence Committee to Ensure the Integrity of the Criminal Process* (2006).

American Bar Association. Criminal Justice Section. *Report to the House of Delegates: Resolution Adopting the American Bar Association Statement of Best Practices for Promoting the Accuracy of Eyewitness Identification Procedures dated August 2004.* http://www.abanet.org/leadership/2004/annual/111c.doc

California Commission on the Fair Administration of Justice. *Report and Recommendations Regarding Eyewitness Identification Procedures* (April 13, 2006). http://www.ccfaj.org/documents/reports/eyewitness/official/eyewitnessidrep.pdf

Illinois Governor's Commission on Capital Punishment Report (April 15, 2002). http://www.idoc.state.il.us/ccp/ccp/reports/commission_report/index.html

National Institute of Justice. *Convicted by Juries, Exonerated by Science: Case Studies in the Use of DNA Evidence to Establish Innocence After Trial* 15 (June 1996). http://www.ojp.usdoj.gov/nij/pubs-sum/161258.htm

National Institute of Justice, *Eyewitness Evidence: A Guide for Law Enforcement* iii (1999). http://www.ojp.usdoj.gov/nij/pubs-sum/178240.htm

New Jersey Department of Law and Public Safety. Office of the Attorney General. *Memorandum Re: Attorney General Guidelines for Preparing and Conducting Photo and Live Lineup Identification Procedures to all County Prosecutors, Police Chiefs, and Law Enforcement Chief Executives* (April 18, 2001). http://www.state.nj.us/lps/dcj/agguide/photoid.pdf

RECEIVED by MSC 4/12/2022 1:07:49 PM

North Carolina Actual Innocence Commission. *Recommendations for Eyewitness Identification* (October 2003). http://www.ncids.org/News%20&%20Updates/Eyewitness%20ID.pdf

Northampton Police Department. "Eyewitness Identification Procedure." *Administration & Operations Manual* O-408: 1-9. http://www.innocenceproject.org/docs/Northampton_MA_ID_Protocols.pdf

Santa Clara County Police Chiefs' Association. Santa Clara County Line-up Protocol. http://www.innocenceproject.org/docs/Santa_Clara_eyewitness.pdf

Virginia State Crime Commission. "Mistaken Eyewitness Identification." *Report to the Governor and the General Assembly of Virginia*, House Document No. 40 (2005). http://leg2.state.va.us/dls/h&sdocs.nsf/By+Year/HD402005/$file/HD40.pdf

Wisconsin Office of the Attorney General. Training and Standards Bureau. *Model Policy and Procedures for Eyewitness Identification* (2005). http://www.doj.state.wi.us/dles/tns/eyewitness.asp

3. Field Studies

Klobuchar, A., and H.L. Caligiuri. "Protecting the Innocent/Convicting the Guilty: Hennepin County's Pilot Project in Blind Sequential Eyewitness Identification." *William Mitchell Law Review* 32 (2005): 1-26

Klobuchar, A., N.K.M. Steblay, and H.L. Caligiuri. "Improving Eyewitness Identifications: Hennepin County's Blind Sequential Lineup Pilot Project." *Cardozo Public Law, Policy & Ethics Journal* 4 (2006): 381-413. http://web.augsburg.edu/~steblay/Improving_Eyewitness_Identifications.pdf

Malpass, R.S. "Notes on the Illinois Pilot Program on Sequential Double-Blind Identification Procedures." *Public Interest Law Reporter* 11, no. 2 (2006): 5-47. http://eyewitness.utep.edu/Documents/Malpass06NotesOnTheIllinoisPilotProgram.pdf

Mecklenburg, S.H. *Report to the Legislature of the State of Illinois: The Illinois Pilot Program on Sequential Double-Blind Identification Procedures* (March 2006). http://www.chicagopolice.org/IL%20Pilot%20on%20Eyewitness%20ID.pdf

Steblay, N. "Observations on the Illinois Lineup Data" (May 3, 2006), http://web.augsburg.edu/~steblay/ObservationsOnTheIllinoisData.pdf.

Wells, G.L. "Gary L. Wells' comments on the Mecklenburg Report" (Updated May 2006), http://www.psychology.iastate.edu/FACULTY/gwells/Illinois_Project_Wells_comments.pdf.

RECEIVED by MSC 4/12/2022 1:07:49 PM

ENDNOTES

[1] *See* Edward Connors et al., Nat'l Inst. of Justice, U.S. Dep't of Justice, *Convicted by Juries, Exonerated by Science: Case Studies in the Use of DNA Evidence to Establish Innocence After Trial* 2, 15 (1996); Gary L. Wells et al., *Eyewitness Identification Procedures: Recommendations for Lineups and Photospreads*, 22 Law & Hum. Behav. 603, 605 (1998); Samuel R. Gross et al., *Exonerations in the United States, 1989 Through 2003*, 95 J. Crim. L. & Criminology 523, 542 (2005).

[2] *See* current DNA exoneration statistics and case profiles at The Innocence Project at the Benjamin N. Cardozo School of Law, http://www.innocenceproject.org.

[3] Gary L. Wells, *Eyewitness Identification: Systemic Reforms*, 2006 Wis. L. Rev. 615, 632 (2006).

[4] Technical Working Group for Eyewitness Evidence, Nat'l Inst. of Justice, U.S. Dep't of Justice, *Eyewitness Evidence: A Guide for Law Enforcement* (1999), http://www.ojp.usdoj.gov/nij/pubs-sum/178240.htm.

[5] Wells, *supra* note 3, at 635.

[6] *Id*. at 623.

[7] Gary L. Wells et al., *Eyewitness Identification Procedures: Recommendations for Lineups and Photospreads*, 22 Law & Hum. Behav. 603, 618 (1998).

[8] *Id*. at 614.

[9] *Id*.

[10] *Id*. at 624.

[11] *Id*.

[12] James M. Doyle, *True Witness: Cops, Courts, Science, and the Battle Against Misidentification* 17 (2005).

[13] *Id*.

[14] Gary L. Wells, *Applied Eyewitness-Testimony Research: System Variables and Estimator Variables*, 36 J. Personality & Soc. Psych. 1546, 1548 (1978).

[15] Wells, *supra* note 3, at 625.

[16] *Id*. *See also* Nancy M. Steblay, *Social Influence in Eyewitness Recall: A Meta-Analytic Review of Lineup Instruction Effects*, 21 Law & Hum. Behav. 283, 284-85 (1997).

[17] Wells et al., *supra* note 7, at 632.

[18] *Id*. at 633.

[19] Wells, *supra* note 3, at 624.

[20] Wells et al., *supra* note 7, at 635.

[21] Wells, *supra* note 3, at 629.

[22] Nancy Steblay et al., *Eyewitness Accuracy Rates in Sequential and Simultaneous Lineup Presentations: A Meta-Analytic Comparison*, 25 Law & Hum. Behav. 459, 464 (2001).

[23] *Id*. at 469.

[24] Wells, *supra* note 3, at 627.

[25] Amy Klobuchar, Nancy M. Steblay & Hilary Lindell Caligiuri, *Improving Eyewitness Identifications: Hennepin County's Blind Sequential Lineup Pilot Project*, 4 Cardozo Pub. L. Pol'y & Ethics J. 381, 410 (2006).

[26] Sheri H. Mecklenburg, *Report to the Legislature of the State of Illinois: The Illinois Pilot Program on Sequential Double-Blind Identification Procedures* (March 2006), http://www.chicagopolice.org/IL%20Pilot%20on%20Eyewitness%20ID.pdf.

[27] *See, e.g.*, Gary L. Wells, Comments on the Mecklenburg Report (Updated May 2006), http://www.psychology.iastate.edu/FACULTY/gwells/Illinois_Project_Wells_comments.pdf; Nancy Steblay, Observations on the Illinois Lineup Data, (May 3, 2006), http://web.augsburg.edu/~steblay/ObservationsOnTheIllinoisData.pdf.

[28] *"Yes, I'm sure-That's The One!"* Memory Loss & the Brain, The Newsletter of the Memory Disorders Project at Rutgers University, Summer 2003, http://www.memorylossonline.com/eyewitness.htm.

[29] *See, e.g.*, State of Wisconsin, Office of the Attorney General, Model Policy and Procedure for Eyewitness Identification (2005) [hereinafter Wisconsin Model Policy], http://www.doj.state.wi.us/dles/tns/EyewitnessPublic.pdf.

[30] State v. Delgado, 188 N.J. 48, 902 A.2d 888 (2006).

[31] North Carolina Actual Innocence Commission, Recommendations for Eyewitness Identification (October 2003), http://www.ncids.org/News%20&%20Updates/Eyewitness%20ID.pdf.

[32] *See* Wisconsin Model Policy, *supra* note 29.

[33] Telephone interview with Ken Hammond, WI Dep't. of Justice Training and Standards Bureau (Jan. 26, 2007).

[34] Klobuchar, Steblay & Caligiuri, *supra* note 25, at 411.

[35] Wells, *supra* note 3, at 642.

[36] *See* Anya Sostek, *Blind Sighted: Eyewitness Identification Doesn't Always Mesh with DNA Evidence, and That's Leading Police Departments To Rethink Their Lineup Procedures*, Governing Magazine, May 2006.

[37] *See* John J. Farmer, Jr., Office of the Attorney General, Attorney General Guidelines for Preparing and Conducting Photo and Live Lineup Identification Procedures (April 18, 2001), http://www.state.nj.us/lps/dcj/agguide/photoid.pdf.

[38] *See* Gina Kolata and Iver Peterson, *New Jersey Is Trying New Way For Witnesses to Say, 'It's Him,'* New York Times, July 21, 2001.

[39] *See* Bill Moushey and Nathan Crabbe, *Questionable Identifications Sent 2 to Jail*, Pittsburgh Post-Gazette, May 9, 2005.

[40] *Id*.

[41] *See Police and Prosecutors Team Up for Better Eyewitness IDs*, November 3, 2003, http://www.psychology.iastate.edu/FACULTY/gwells/Minnesota_new_procedures.html.

[42] *See* Jennifer Thompson, *'I Was Certain, but I Was Wrong,'* New York Times, June 18, 2000.

[43] *See Local Police Unaware of Federal Eyewitness Guidelines*, Associated Press, May 9, 2005.

RECEIVED by MSC 4/12/2022 1:07:49 PM

The Justice Project is comprised of two nonpartisan organizations dedicated to fighting injustice and to creating a more humane and just world. The Justice Project, Inc., which lobbies for reform, and The Justice Project Education Fund, which increases public awareness of needed reforms, work together on the Campaign for Criminal Justice Reform to reaffirm America's core commitment to fairness and accuracy by designing and implementing national and state-based campaigns to advance reforms that address significant flaws in the American criminal justice system, with particular focus on the capital punishment system.

This report is made possible primarily through a grant from The Pew Charitable Trusts to The Justice Project Education Fund. The opinions expressed are those of the author(s) and do not necessarily reflect the views of the Trusts. For additional information, questions or comments, please contact our offices at (202) 638-5855, or email *info@thejusticeproject.org*.

THE JUSTICE PROJECT

1025 Vermont Avenue, NW • Third Floor • Washington, DC 20005
202 638-5855 • Fax 202 638-6056 • www.TheJusticeProject.org

RECEIVED by MSC 4/12/2022 1:07:49 PM

APPENDIX E

Eyewitness Memory Still Not Common-Sense Table

RECEIVED by MSC 4/12/2022 1:07:49 PM

Factors Found to Impact Eyewitness Accuracy	*Factor Descriptions*	% Experts (n = 64)	% Jurors (n = 111)	χ^2 (df = 1) and significance (p)
Accuracy and confidence	An eyewitness's confidence is not a good predictor of his or her identification accuracy.	87	38	38.02 (0.000)
Alcoholic intoxication	Alcoholic intoxication impairs an eyewitness's later ability to recall persons and events.	90	96	1.64 (*ns*)
Attitudes and expectations	An eyewitness's perception and memory for an event may be affected by his or her attitudes and expectations.	92	81	3.96 (0.047)
Child suggestibility	Young children are more vulnerable than adults to interviewer suggestion, peer pressures and other social influences.	94	82	4.75 (0.029)
Child witness accuracy	Young children are less accurate as witnesses than are adults.	70	32	24.60 (0.000)
Colour perception	Judgements of colour made under monochromatic light (e.g. an orange street light) are highly unreliable.	63	33	14.02 (0.000)
Confidence malleability	An eyewitness's confidence can be influenced by factors that are unrelated to identification accuracy.	95	50	38.04 (0.000)
Cross-race bias	Eyewitnesses are more accurate when identifying members of their own race than members of other races.	90	47	33.32 (0.000)
Description matching	The more members of a lineup resemble a witness's description of the culprit, the more accurate an identification of the suspect is likely to be.	71	35	21.94 (0.000)
Discriminability	It is possible to reliably differentiate between true and false memories.	32	47	4.08 (0.043)
Elderly witness	Elderly eyewitnesses are less accurate than are younger adults.	50	23	12.94 (0.000)
Event violence	Eyewitnesses have more difficulty remembering violent than non- violent events.	36	26	1.87 (*ns*)
Exposure time	The less time an eyewitness has to observe an event, the less well he or she will remember it.	81	47	19.93 (0.000)
False childhood memories	Memories people recover from their own childhood are often false or distorted in some way.	68	35	18.40 (0.000)
Forgetting curve	The rate of memory loss for an event is greatest right after the event and then levels off over time.	83	33	39.79 (0.000)
Hypnotic accuracy	Hypnosis increases the accuracy of an eyewitness's reported memory.	45	19	13.86 (0.000)
Hypnotic suggestibility	Hypnosis increases suggestibility to leading and misleading questions.	91	24	71.44 (0.000)
Identification speed	The more quickly a witness makes an identification upon seeing the lineup, the more accurate he or she is likely to be.	40	65	9.68 (0.002)
Lineup fairness	The more members of a lineup resemble the suspect, the higher is the likelihood that identification of the suspect is accurate.	70	32	24.60 (0.000)
Lineup instructions	Police instructions can affect an eyewitness's willingness to make an identification.	98	41	57.59 (0.000)
Long-term repression	Traumatic experiences can be repressed for many years and then recovered.	22	73	42.71 (0.000)
Mug shot-induced bias	Exposure to mug shots of a suspect increases the likelihood that the witness will later choose that suspect in a lineup.	95	59	27.20 (0.000)
Post event information	Eyewitness testimony about an event often reflects not only what they actually saw but also information they obtained later on.	94	60	23.68 (0.000)
Presentation format	Witnesses are more likely to misidentify someone by making a relative judgement when presented with a simultaneous (all at the same time) as opposed to sequential (one at a time) lineup.	81	31	41.62 (0.000)
Show-ups	The use of a one-person show-up instead of a full lineup increases the risk of misidentification.	74	51	8.21 (0.004)
Stress	Very high levels of stress impair the accuracy of eyewitness testimony.	60	68	1.91 (*ns*)
Trained observers	Police officers and other trained observers are no more accurate as eyewitnesses than is the average person.	39	28	2.98 (*ns*)
Unconscious transference	Eyewitnesses sometimes identify as a culprit someone they have seen in another situation or context.	81	30	43.14 (0.000)
Weapon focus	The presence of a weapon impairs an eyewitness's ability to accurately identify the perpetrator's face.	87	39	36.70 (0.000)
Wording of questions	An eyewitness's testimony about an event can be affected by how the questions put to that witness are worded.	98	85	8.32 (0.004)

Source: Benton, T. R., Ross, D. F., Bradshaw, E., Thomas, W. N., & Bradshaw, G. S. (2006). Eyewitness memory is still not common sense: Comparing jurors, judges and law enforcement to eyewitness experts. *Applied Cognitive Psychologyn, 20(1),* 115-129.

RECEIVED by MSC 4/12/2022 1:07:49 PM

APPENDIX F

Comprehensive Evaluation of Show Ups

A Comprehensive Evaluation of Showups

Jeffrey S. Neuschatz, Stacy A. Wetmore, Kylie N. Key, Daniella K. Cash, Scott D. Gronlund, and Charles A. Goodsell

A showup is a procedure in which police present an eyewitness with a single person, either live or in a photograph, for the purpose of identification (Valentine, Davis, Memon, & Roberts, 2012). Typically, these one-on-one confrontations occur in the field, and in close spatial and temporal proximity to the crime (Behrman & Davey, 2001). Showup identifications are an alternative to lineup identifications (which involve multiple individuals being presented to the witness); however, the U.S. Supreme Court, state courts, and social science researchers have stated repeatedly that showup identifications are less reliable than lineup identifications (Gronlund et al., 2012; *Stovall v. Denno*, 1967; Wetmore et al., 2015a). A study by Gronlund et al. (2012) found that showups led to worse identification performance than lineups. This finding has been confirmed by several other researchers (Clark & Godfrey, 2009; Steblay, Dysart, Fulero, & Lindsay, 2003; Wetmore et al., 2015a). Garrett (2011) reviewed trial transcripts from 160 DNA exoneration cases and found that 34 % (53/160) involved misidentifications from showups. Nevertheless, showups remain one of the most widely employed identification procedures (see Goodsell, Wetmore, Gronlund, and Neuschatz, 2013). The estimates of showup identifications as a percentage of all identifications range from 30 to 77 % (Gonzalez, Ellsworth, & Pembroke, 1993; McQuiston & Malpass, 2001). Therefore, showups are a significant percentage of all identifications and warrant further investigation.

This chapter examines showups and lineups, particularly the comparative claim that showups are inherently more suggestive than lineups. The chapter also examines

J.S. Neuschatz (✉) • K.N. Key • D.K. Cash
The University of Alabama in Huntsville, Huntsville, AL, USA
e-mail: neuschaj@uah.edu

S.A. Wetmore • S.D. Gronlund
University of Oklahoma, Norman, OK, USA

C.A. Goodsell
Canisius College, Buffalo, NY, USA

© Springer International Publishing Switzerland 2016

M.K. Miller, B.H. Bornstein (eds.), *Advances in Psychology and Law*,
Advances in Psychology and Law 1, DOI 10.1007/978-3-319-29406-3_2

if there are situations in which a showup is comparable in performance to a lineup. In the sections that follow, we operationally define terms important to the comparison of these identification procedures, as well as outline advantages and disadvantages of showups relative to lineups. The extant literature is meta-analyzed comparing showups and lineups, using traditional measures (e.g., correct identification rates, false identification rates, and ratio-based probative value measures). These analyses are followed by Receiver Operating Characteristic (ROC) analyses, where feasible. Finally, the confidence–accuracy relationship yielded by each identification procedure is assessed using both point-biserial correlations and confidence calibration. We conclude by addressing proposed theoretical explanations, real-world implications of the findings, and ways in which showup procedures might be improved to yield more accurate identifications.

Advantages and Disadvantages of Showups

In order to understand the advantages and disadvantages of showups, it is important to define some terms that will be used throughout the chapter. Choices made from both perpetrator present (PP) and perpetrator absent (PA) identification procedures are compared when trying to determine which procedure is superior. Perpetrator present lineups (or showups) contain the actual perpetrator; perpetrator absent lineups (or showups) replace the perpetrator with an innocent suspect designated by the researcher. This is only true in the lab (in the field, the suspect is not known to be innocent). There are two correct choices that can be made: correct identifications (i.e., identifications of the perpetrator) in a perpetrator present identification procedure or correctly indicating that the perpetrator is not there from a perpetrator absent procedure. There are three possible errors that can occur: false identifications, incorrect rejections, and filler identifications (lineups only). False identifications refer to identifications of innocent suspects; incorrect rejections arise when a rejection is made from a perpetrator present lineup (or showup); filler identifications refer to identifications of people in the lineup who are known to be innocent. Because fillers are known to be innocent, filler identifications are not considered harmful errors. If a filler is chosen, the lineup administrator knows that the witness is unsure of the perpetrator's identity and should not be trusted as an accurate eyewitness. We focus on three identification responses for purposes of evaluating identification performance: correct identifications from perpetrator present identification procedures, false identifications from perpetrator absent identification procedures, and the confidence with which those decisions are made.

Showup identifications may confer performance advantages over traditional lineups. First, showups have the potential to be conducted faster than lineups, and therefore implicate criminals or absolve innocent people of suspicion quickly. Second, as has been well documented, memory performance decreases with time, so it is better to test memory (e.g., administer an eyewitness identification test) at a short rather than long delay (e.g., Light, 1996). Consistent with basic memory research, research

on eyewitness identifications has shown that witnesses tend to perform more poorly at identifying guilty suspects as time progresses (Clark & Godfrey, 2009). Consistent with the psychological research, the U.S. Supreme Court and state courts have held that one of the circumstances under which it is acceptable to conduct a showup is when it occurs close in time to the incident in question (see Gardner & Anderson, 2004; *Stovall v. Denno*, 1967). Thus, a potential advantage of showups is that they may result in better memory performance than lineups due to a shorter delay—retention interval—between the incident and the administration of the identification procedure. Given the shorter delay, the perpetrator may not have time to alter distinctive features noticed by the witness (e.g., clothing, facial hair). Thus, the speed at which a showup can be conducted should afford extra cues like clothing match that are likely not available in a lineup conducted after a delay. These extra cues could aid recognition memory.

Despite these potential advantages, the U.S. Supreme Court repeatedly has argued against the use of showups (*Stovall v. Denno*, 1967; *United States v. Wade*, 1967). In *Stovall*, the court argued that showing suspects singly for the purpose of identification was a widely condemned practice. Consequently, showups are widely thought to be suggestive and should, with few exceptions, not be used by law enforcement. From an applied research perspective, the disadvantages arise from three factors: administration, lack of fillers, and choosing rates. We discuss these next.

First, showups are difficult to administer in a double-blind manner. Double-blind administration of a lineup refers to a situation where both the lineup administrator and the eyewitness are unaware of the identity of the suspect before the identification is conducted. This is problematic because an investigator who knows who the suspect is can intentionally or inadvertently convey the "correct" answer to the eyewitness and therefore influence the identification decision (see Haw & Fisher, 2004), without the witness even being aware of the influence (Garrioch & Brimacombe, 2001). When evaluating a positive identification under these circumstances, it is impossible to know whether the identification is the product of the witness's memory for the perpetrator, influenced by the investigator, or some combination of the two. Given that the purpose of an identification is to test the hypothesis that the suspect is the perpetrator, ruling out alternative explanations for positive identifications is critical. Haw and Fisher (2004) minimized witness contact with the lineup administrator, which resulted in decreased false identifications without reducing hits. It is important to note that blind administration has never, to our knowledge, been tested with showups; however, given the long history demonstrating experimenter bias in the psychology literature, it is reasonable to assume that the same concerns apply to showups (see Rosenthal, 1966; Rosenthal & Jacobson, 1968).

Second, there are no fillers in a showup, which means there is no protection against witnesses who are guessing (i.e., willing to respond despite very low confidence). The problem is compounded by the fact that the likelihood of choosing the suspect in showups is greater than in lineups ($1/n$ where n is greater than 2 in a lineup). Luus and Wells (1991) outlined some of the important functions of fillers in

a well-constructed, fair lineup. A selection by the witness of a filler is a known error, which gives the lineup administrator information regarding the accuracy of the eyewitness. The filler also serves as a control for guessing—the suspect should not be chosen more often than each of the fillers if a witness has no memory of the perpetrator (Malpass & Lindsay, 1999).

Finally, a witness may be more likely to make an identification from a showup regardless of whether the perpetrator is present or not (see Dysart & Lindsay, 2007; Goodsell et al., 2013). It is commonly assumed that showups will produce higher choosing rates because they include only one individual—witnesses are more likely to choose if there is only one option (Kassin, Tubb, Hosch, & Memon, 2001). According to signal detection theory (Green & Swets, 1966), participants use a more liberal criterion when there is only one option. The higher choosing rate will lead to more correct identifications but also more false identifications of suspects. This translates into a greater potential of innocent people being identified, indicted, and potentially incarcerated from showups than lineups. In fact, in the Steblay et al. (2003) meta-analysis comparing showups and lineups, participants did have higher choosing rates in showups than lineups. Consistent with this meta-analysis, Meissner, Tredoux, Parker, and MacLin (2005) examined face memory utilizing identification procedures ranging in size from 1 to 12 options and found that criterion became more liberal as lineup size decreased, until lineup size equaled 1 (showup).

Regarding choosing rates, laboratory studies probably underestimate the actual choosing rates in showups because researchers cannot simulate the social pressure that a real eyewitness feels (Dysart, Lindsay, & Dupuis, 2006). Indeed, in their archival analysis of real-world cases, Behrman and Davey (2001) reported that the suspect identification rate in showups was 76 % as opposed to 48 % in lineups. According to the Best Practice Guidelines outlined by the Technical Working Group for Eyewitness Evidence (1999), using unbiased instructions, or indicating that the perpetrator may or may not be present in the identification procedure, may alleviate this pressure to choose. Typically, lineup administrators employ unbiased instructions, but this likely is less true for showups. Given the different circumstances surrounding a showup, unbiased instructions may be less effective. Although the use of unbiased instructions has never been examined using showup identification procedures, research has demonstrated that the expectation that a perpetrator is present in a lineup eliminates the protection of unbiased instructions (Quinlivan et al., 2011). There are several reasons to believe that the expectation of perpetrator presence are even greater in a showup than a lineup due to circumstances surrounding showups (proximity, time, one person).

Showups Versus Lineups

The remainder of the chapter is organized around the central question: Are showups always a less reliable identification procedure than lineups? To tackle this question, in the subsections that follow, we compare showups and lineups across a

range of factors. However, we begin by collapsing over all conditions and then break down the data into conditions that are forensically or theoretically important to the question.

First, we compare showups to simultaneous (all photographs shown at the same time) and sequential (photographs shown one at a time) lineups. There has been a vigorous debate in the literature about which lineup yields better accuracy (for a review, see Clark, 2012; Gronlund, Andersen, & Perry, 2013; Steblay, Dysart, & Wells, 2011); therefore, we felt it is necessary to compare both simultaneous and sequential lineups to showups. Next, we compare showups to simultaneous and sequential lineups with suspects in different lineup positions (2 or 5). Previous research has identified suspect position as a factor that moderates the effectiveness of sequential lineups (Carlson, Gronlund, & Clark, 2008; Gronlund, Carlson, Dailey, & Goodsell, 2009). Gronlund et al. (2012) found that when the suspect appears early in the sequential lineup (position 2), performance was no better than a showup, but when the suspect appears late in the sequential lineup (position 5), sequential performance was as good as simultaneous performance. Finally, we examine variables that should supposedly provide an advantage for showups (i.e., short retention interval, similar cues at encoding and test). We also compared showups to fair and biased lineups; fair lineups contain fillers that all match the general description of a suspect, whereas biased lineups contain only one or two plausible options. Because of the lack of viable options in a biased lineup, a showup might be preferred to a biased lineup. We examined these comparisons using traditional meta-analysis, ROC analysis, point-biserial correlations, and confidence calibration.

Meta-analysis

Steblay et al. (2003) conducted the first meta-analytic comparison of showups and lineups. They analyzed eight published articles that included 12 tests of identification performance in showups and lineups, and found that the choosing rate, collapsed over perpetrator present and perpetrator absent conditions, was significantly higher in lineups (54 %) than in showups (27 %). Despite choosing less often, the laboratory data indicated that showup choices were more accurate: Correct decisions (perpetrator identifications from perpetrator present conditions + correct rejections from perpetrator absent conditions) were significantly higher in showups (69 %) than in lineups (51 %). In addition, the number of incorrect identifications from perpetrator absent conditions was significantly lower in showups (15 %) than in lineups (43 %). However, Steblay et al. (2003) argued that this was a misleading comparison because filler identifications from lineups were not "dangerous errors." Instead, they argued that it made more sense to compare "dangerous errors" involving only the identifications of an innocent suspect. This is important because researchers must compare perpetrator present and perpetrator absent conditions to get an understanding of the utility of an identification procedure—comparing the benefits of identifying the guilty to the costs of accusing the innocent.

Only five of the 12 tests reviewed by Steblay et al. (2003) included a designated innocent suspect. When they focused on just those studies (Dekle, Beal, Elliott, & Huneycutt, 1996; Yarmey, Yarmey, & Yarmey, 1994, 1996), they found that the false identification rate from showups (23 %) was higher than for lineups (10 %). But unfortunately, Steblay et al. failed to consider correct identifications. In the aforementioned studies the correct identification rate for showups (59 %) was greater than for lineups (39 %). Thus, the benefit of fewer false identifications was offset by the cost of fewer correct identifications. This complicates the decision about which identification procedure is best (see Clark (2012) for a discussion of the cost-benefit tradeoff regarding showups and lineups).

A second meta-analysis was conducted by Clark and Godfrey (2009). They included the five studies from Steblay et al. (2003), plus three additional studies, which resulted in a total of 15 showup–lineup comparisons (all with adult participants). Contrary to Steblay et al., Clark and Godfrey found that correct identification rates in perpetrator present conditions and innocent suspect identifications from perpetrator absent conditions were not significantly different between lineups and showups. Showups also were not found to significantly increase choosing rates relative to lineups.

Clark and Godfrey (2009), however, argued that focusing on correct identifications in perpetrator present and correct rejections in perpetrator absent lineups places lineups at a disadvantage. Witnesses can choose a filler in a lineup but not a showup, and every filler choice reduces the correct rejection rate of a lineup. Clark and Godfrey argued that conditional probability (CP = correct identification/[correct identifications + false identifications]) was a better measure of identification performance. Clark and Godfrey actually reported 1-CP or what they termed the innocence risk. The innocence risk for lineups (.21) was significantly less than for showups (.31), indicating that showups put an innocent suspect at greater risk of being falsely identified. However, this result was mitigated by retention interval. There was no difference at short retention intervals, but at 2- and 24-h delays, the innocence risk for showups was much higher than for lineups.

Updated Meta-analysis

We used the studies examined by the Clark and Godfrey (2009) meta-analysis, and studies published since. To be included, studies had to compare showups and lineups and use adult participants. When no designated innocent suspect was indicated, the false identification rates were estimated by using the probability of selecting any lineup member (1/number of lineup members).[1] There were a total of 21 showup–lineup comparisons, and we utilized the method of analysis in Clark and Godfrey (2009): *t*-tests to compare identification procedures and Cohen's *h* as a measure of effect size (see Cohen, 1988). The experiment, not the participants, is the unit of

[1] Although Lawson and Dysart (2014) included showups and lineups in their paper, they did not include correct identification rates. Attempts to get that information were unsuccessful.

analysis. The addition of these extra studies (compared to 13 in Clark and Godfrey and 8 in Steblay et al., 2003) also allowed us to conduct ROC analyses, which will be reported in the next section. Table 1 contains the correct identification rates, false identifications rates, and probative values for all the studies included in the analyses.[2]

As can be seen in Table 1, the correct and false identification rates for showups and lineups were similar and not significantly different. Consistent with Clark and Godfrey (2009), the only significant differences were observed for the probative value measure. This indicates that, given that the witness makes an identification, it is more likely that the witness will choose the perpetrator from a lineup than from a showup. Stated more simply, showups put innocent suspects at greater risk than lineups. In addition to overall correct and false identification rates, we compared showups conducted immediately to lineups conducted after a delay (defined as a retention interval greater than 24 h).[3] Because there are very few delayed lineups ($n = 5$), we only present the descriptive statistics. Surprisingly, even with a retention interval greater than 24 h, there are more correct identifications ($M = .52$) and fewer false identifications ($M = .13$) from lineups than showups conducted at shorter retention intervals (Ms = .45 and .18 for the correct and false identification rates, respectively). Thus, although common sense might lead one to expect witnesses to display a showup advantage at a short retention interval, the data contradict this supposition.

However, before we accept any of these conclusions, we have to acknowledge the problems with using correct/false identification rates and ratio-based probative value measures to determine which identification procedure is superior. These measures of performance are confounded by response bias (different choosing rates across procedures). Thus, a measure like ROC analysis that can disentangle response bias from discriminability must be used (see Gronlund & Neuschatz, 2014; Wixted & Mickes, 2012).

ROC Analysis

As Gronlund and colleagues argue (Gronlund & Neuschatz, 2014; Gronlund, Wixted, & Mickes, 2014), ROC analysis can determine if an identification procedure results in a performance benefit (i.e., better discriminability of the perpetrator from the innocent suspect), a difference in response bias, or both. ROC analysis is grounded in signal-detection theory (Macmillan & Creelman, 2005) and is a well-established analytic technique used for evaluating diagnostic decisions in

[2] One measure of probative value (PV) is the probability of choosing the guilty suspect given that a suspect (innocent or guilty) was chosen (=[(guilty suspect identifications from perpetrator present lineups)/(guilty suspect identifications from perpetrator-present lineups + innocent suspect identifications from perpetrator-absent lineups)]).

[3] Given the way the data were reported, it was impossible to analyze other variables, but we will examine those in the next section with ROC analyses.

Table 1 Proportion of correct and false identifications and conditional probabilities for studies included in the meta-analysis

Study	Lineup size	Delay	Lineup			Showup			Cohen's H		
			CID	FID	PV	CID	FID	PV	CID	FID	PV
Wagenaar and Veefkind (1992)[a]	2	0 h	.56	.12	.82	.35	.11	.76	.42	.03	.15
Wagenaar and Veefkind (1992)[a]	6	0 h	.5	.07	.88	.35	.11	.76	.30	−.14	.28
Wagenaar and Veefkind (1992)[a]	10	0 h	.42	.05	.89	.35	.11	.76	.14	−.23	.38
Wagenaar and Veefkind (1992) Exp. 2[a]	6	7 days	.75	.05	.94	.5	.15	.77	.52	−.34	.51
Gonzalez et al. (1993) Exp. 1	6	10 min	.64	.5	.56	.21	.07	.75	.90	1.04	.20
Gonzalez et al. (1993) Exp. 2	6	0 h	.25	.2	.56	0	.03	0.00	1.05	.61	2.09
Dekle et al. (1996)[a]	6	0 h	.3	.06	.83	.28	.04	.88	.04	.09	−.12
Dekle (1997)	6	2–3 days	.3	.06	.83	.35	.05	.88	−.11	.04	−.14
Dekle (1997)	6	7 days	.31	.1	.76	.3	.13	.70	.02	−.09	.16
Yarmey et al. (1994)[a]	6	5 min	.46	.05	.90	.57	.12	.83	−.22	−.26	.21
Yarmey et al. (1996)	6	0 h	.49	.16	.75	.7	.18	.80	−.43	−.05	−.12
Yarmey et al. (1996)	6	5 h	.39	.33	.54	.64	.44	.59	−.51	−.23	−.12
Yarmey et al. (1996)	6	2 h	.36	.14	.72	.54	.58	.48	−.36	−.96	.50
Yarmey et al. (1996)	6	1 day	.32	.14	.70	.55	.53	.51	−.47	−.86	.39
Lindsay, Pozzulo, Craig, Lee, and Corber (1997) simultaneous	6	0 h	.55	.06	.91	.5	.07	.88	.10	−.04	.10
Lindsay et al. (1997) sequential	6	0 h	.62	.04	.94	.5	.07	.88	.24	−.13	.21
Gronlund et al. (2012)	6	10 min	.65	.18	.78	.49	.24	.67	.32	.15	.32

(continued)

Table 1 (continued)

Study	Lineup size	Delay	Lineup			Showup			Cohen's H		
			CID	FID	PV	CID	FID	PV	CID	FID	PV
Valentine et al. (2012)	6	15 min	.68	.12	.85	.65	.05	.93	.06	.26	−.25
Wetmore et al. (2015a)	6	0 h	.74	.28	.73	.62	.44	.58	.26	−.34	.30
Wetmore et al. (2015b)	6	2 days	.72	.3	.71	.56	.41	.58	.34	−.23	.27
Key et al. (2015)	6	0 h	.65	.18	.78	.44	.26	.63	.42	−.20	.34
M			.51	.15	.78	.45	.20	.70	.14	−.09	.27
SD									.41	.42	.47
t(20)									1.89	−.97	2.62

Note. Bold studies were included in Clark and Godfrey (2009)
[a]Studies were included in Steblay et al. (2003). Wetmore et al. (2015b) was included in the data analyses but not in the table because they only used showups. PV = CID/(CID+FID)

many different domains (Swets, Dawes, & Monahan, 2000). An eyewitness ROC curve is created by plotting the correct identification rate against false identification rate at each level of witness confidence (see Fig. 1 for an example). Filler identifications are excluded, just as they are from probative value calculations, because the identification of a filler is not a harmful error. The lower left-hand point on the ROC curve depicts the most confident correct and false identification responses (confidence level of 7 or the most conservative response bias). The next point on the ROC curve depicts the most and second-most confident correct and false identifications (confidence of 7 and 6), and so forth for the remainder of the confidence scale. Thus, ROCs depict a cumulative record of suspect identification decisions across the entire confidence scale. The data are summarized by a trend-line to better depict each curve.

The identification procedure whose ROC curve is closest to the upper left corner of the space exhibits better discriminability than ROCs that fall below it. Discriminability can be assessed by computing the area under the curve (AUC) for each identification procedure; the greater the AUC, the better the discriminability of that procedure. Lineup ROCs are truncated and do not trace out over the entire probability space (0–1) because the curve only extends to the highest false identification rate (recall that filler identifications are excluded). Therefore, the partial area under the curve (pAUC)[4] must be computed.

[4]pAUC values are computed using a false ID rate range from 0 to q, where q is set to a value slightly greater than the maximum false ID rate for the ROCs used in a comparison (see Wixted & Mickes, 2012). To evaluate different identification procedures, statistically compare the pAUCs and report the pAUCs, D, and p values. D is defined as (AUC1 − AUC2)/s, where s is the standard error of the difference between the two pAUCs (Robin et al., 2011). The standard error is estimated by the bootstrap method using 10,000 bootstraps (see Mickes, Flowe, & Wixted, 2012; for a tutorial, see Gronlund et al., 2014).

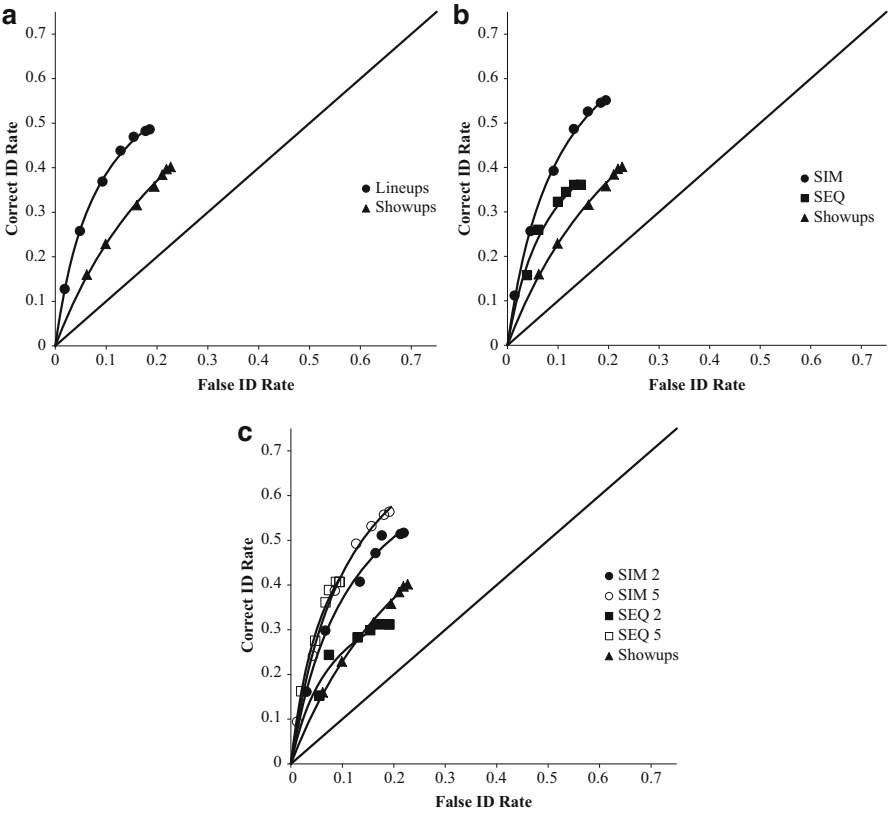


Fig. 1 ROC curves comparing lineups and showups. (**a**) Compares lineup and showups overall, (**b**) compares showups versus simultaneous and sequential lineups. (**c**) Compares showups versus lineups separated by lineup type and suspect position

In order to conduct ROC analyses, one needs to have confidence ratings associated with each identification response. This limited the sample of available studies because most researchers do not report confidence ratings in this way.[5] There were six such studies, which included a total of 7877 participants.[6] These studies are italicized in Table 1.

As we did for the evaluation of the data meta-analytically, we first review the ROC data by evaluating showups versus lineups overall, and then consider other variables pertinent to discriminability (e.g., lineup type, suspect position). ROC curves comparing showups and lineups, with showups and lineups broken down as

[5] Gronlund et al. (2012) utilized the same data reported in Gronlund et al. (2009). Wetmore et al. (2015a) had different data but utilized the same stimuli as Gronlund et al. (2009).

[6] A total of 1050 participants in the innocent strong condition were removed from the analyses because performance was poor. Clark (2012) also excluded these data from his meta-analysis.

a function of lineup type, and as a function of suspect position, are all presented in Fig. 1. Panel a shows that participants are better able to discriminate perpetrators from innocent suspects in lineups (pAUC = .08) than showups (pAUC = .05), $D = 7.59$, $p < .001$. Panel b shows that lineups, irrespective of presentation method, result in better performance than showups. The pAUC for the simultaneous lineup (.09) was significantly greater than that of the showup (.05), $D = 8.32$, $p < .001$. Similarly, the pAUC was greater for the sequential lineup (.07) than the showup, $D = 2.24$, $p = .02$. These findings are consistent with expert opinion that showups are less reliable than lineups (Kassin et al., 2001) and confirm what has been reported in previous studies (Gronlund et al., 2012; Wetmore et al., 2015a). Additionally, there was a significant difference between the pAUCs for the simultaneous versus sequential lineup, $D = 3.57$, $p < .001$. This finding is consistent with a number of studies that have recently found better discriminability from simultaneous than sequential lineups (Carlson & Carlson, 2014; Dobolyi & Dodson, 2013; Mickes et al., 2012).

Panel c of Fig. 1 shows the ROC curves for showups and lineups broken down by suspect position. Consistent with Gronlund et al. (2012), we found no difference in discriminability between showups and sequential lineups when the suspect was in position 2. The pAUCs for sequential position 2 (.05) and showups (.05) were not different, $D = .08$, $p = .93$. By contrast, the simultaneous lineup position 2 (pAUC = .08) and the showup were significantly different, $D = 2.80$, $p = .01$. Moreover, the remaining pAUCs were greater than that of the sequential position 2, and therefore significantly superior to showups, Ds > 3.03, $p < .001$, but not different from one another. In sum, simultaneous lineups are better than showups, as are sequential lineups if the suspect occurs late in the lineup.

Thus far, the analyses portray a dim view of showup identifications. However, they fail to consider two situations in which showups might be more reliable than lineups. First, as discussed previously, an immediate showup should confer a memorial advantage relative to a delayed lineup. Second, a showup may induce better discriminability when the clothing of the suspect matches what the perpetrator was wearing. According to the principle of encoding specificity (Tulving & Thomson, 1973), similar cues at study and test (same clothing) should enhance discriminability relative to different cues (different clothing). Moreover, a clothing match should be more likely in a showup because it generally takes place shortly after the crime and in close temporal and spatial proximity (see Valentine et al., 2012), making it less likely that a perpetrator could change clothing before the identification. Although some authors have reported that a clothing match increases false identifications and has no effect on correct identifications (Dysart Lindsay, & Dupuis, 2006; Yarmey et al., 1996), Wetmore, Neuschatz, Gronlund, Key, and Goodsell (2015b) recently found, using ROC analysis, that clothing match enhanced discriminability from showups relative to non-matching clothing. Therefore, it is possible that a clothing match might be beneficial to a showup identification.

To test the retention interval predictions, immediate showups were compared to simultaneous lineups conducted after a 2-day retention interval. This dataset did not include any sequential lineups conducted at a delay. Surprisingly, as can be seen

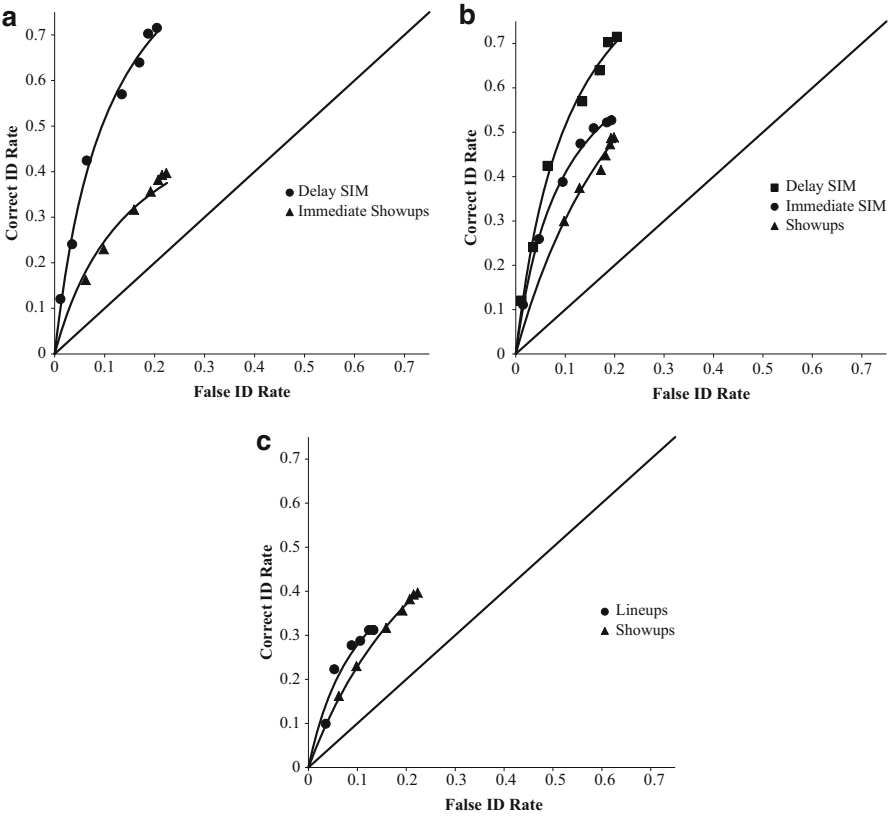


Fig. 2 ROC curves comparing simultaneous lineups and showups. (**a**) Depicts ROCs comparing immediate showups versus delayed simultaneous lineups. (**b**) Depicts immediate showups, match condition versus immediate and delay simultaneous lineups, no match. (**c**) Compares fair and biased lineups to showup conditions

in Panel a of Fig. 2, discrimination was significantly worse in showups conducted immediately (pAUC = .05) compared to simultaneous lineups after a 2-day delay (pAUC = .11), $D = 4.68$, $p < .001$.

To test the clothing match predictions, we compared immediate showups in which the clothing was the same during the crime and the identification (the match conditions from Wetmore et al., 2015b) to lineups that were conducted immediately or after a delay with no clothing match. As can be seen in Panel b of Fig. 2, discriminability was significantly better for simultaneous lineups conducted immediately (pAUC = .09) than for showups conducted immediately, (pAUC = .07), $D = 2.72$, $p < .001$, even when the showup suspects wore the same clothing during the crime and identification. Even the delayed simultaneous lineup (pAUC = .11) outperformed the immediate showup with clothing match, $D = 3.15$, $p < .001$.

It is important to note that there were other differences between the clothing match showups (Wetmore et al., 2015b) and the lineups in Wetmore et al. (2015a),

including the use of video showups that showed the suspect's entire body, compared to lineup conditions that utilized photo headshots. However, because the showup revealed the suspect's entire body, it likely provided additional cues to help participants discriminate the guilty from the innocent suspect. Despite this, participants' discrimination ability was still inferior in the showup conditions.

Lineup fairness is another factor that could differentially affect eyewitness performance as a function of lineup procedure (Carlson et al., 2008; Gronlund et al., 2009). Although Wells and Quinlivan (2009) suggested that showups may put an innocent suspect at greater risk than a fair lineup, the showup may be better than placing an innocent suspect in a biased lineup. Lineup fairness refers to the degree to which the fillers in the lineup match the perpetrator. Traditionally, lineup fairness is assessed using Tredoux's *E,* which assesses the nominal size of the lineup—how many viable lineup members there are in the lineup. If a six-person lineup has a Tredoux's *E* of 4, it indicates that there are four viable alternatives from which the witness might choose. For the studies that manipulated lineup fairness (Gronlund et al., 2009; Wetmore et al., 2015a), a fair lineup required a Tredoux's *E* value greater than 4 and a biased lineup required a value near 1. Examples of the fair and biased lineup for the innocent suspect can be seen in Fig. 3.

The ROC curves for the fair lineup, biased lineup, and showup are shown in Fig. 2c. The fair lineup pAUC (.11) and biased lineup pAUC (.12) were significantly greater than the showup, pAUC (.07), $D = 6.59$, $p < .001$ and $D = 7.13$, $p < .001$, respectively. Surprisingly, there was no significant difference between the fair and biased lineup pAUCs, $D = -.95$, $p = .34$, indicating a similar ability to discriminate guilty from innocent suspects irrespective of lineup fairness. However, it would be wrong to conclude that the lack of a discriminability difference signals that our fairness manipulation was too weak because lineup fairness greatly impacted the range over which the curves extend.

The reason the ROC curve for the biased lineup extended over a greater range than the ROC curve for the fair lineup is apparent upon examination of Fig. 4. Figure 4 depicts possible underlying memory strength distributions portraying biased and fair lineups. The distribution in blue represents the range of different memory strengths for the perpetrator that different participants possess. That is, a few participants have an excellent memory for the perpetrator, most have a moderate memory, and a few have a poor memory for the perpetrator. Likewise, the distribution in green represents the varying degrees to which an innocent suspect matches participants' memory for the perpetrator. If a lineup is fair, the fillers should not stand out from an innocent suspect and therefore we can assume that the green distribution also represents the range of memory strengths of these fillers. However, if the lineup is biased it means that fillers are worse matches to the perpetrator. This poor filler distribution (in red) is shifted lower (has lower memory strength) than the distribution of good fillers (in green). Because ROC discriminability is a function of the overlap between the perpetrator and innocent suspect distributions, it is little affected by the distribution of biased fillers. Hence, there is no difference in discriminability as a function of lineup fairness. But the extent of the ROC curve arises from the placement of the confidence criteria (the vertical lines in the figure). The greater range over which the ROC curve for the biased lineup extends is due to

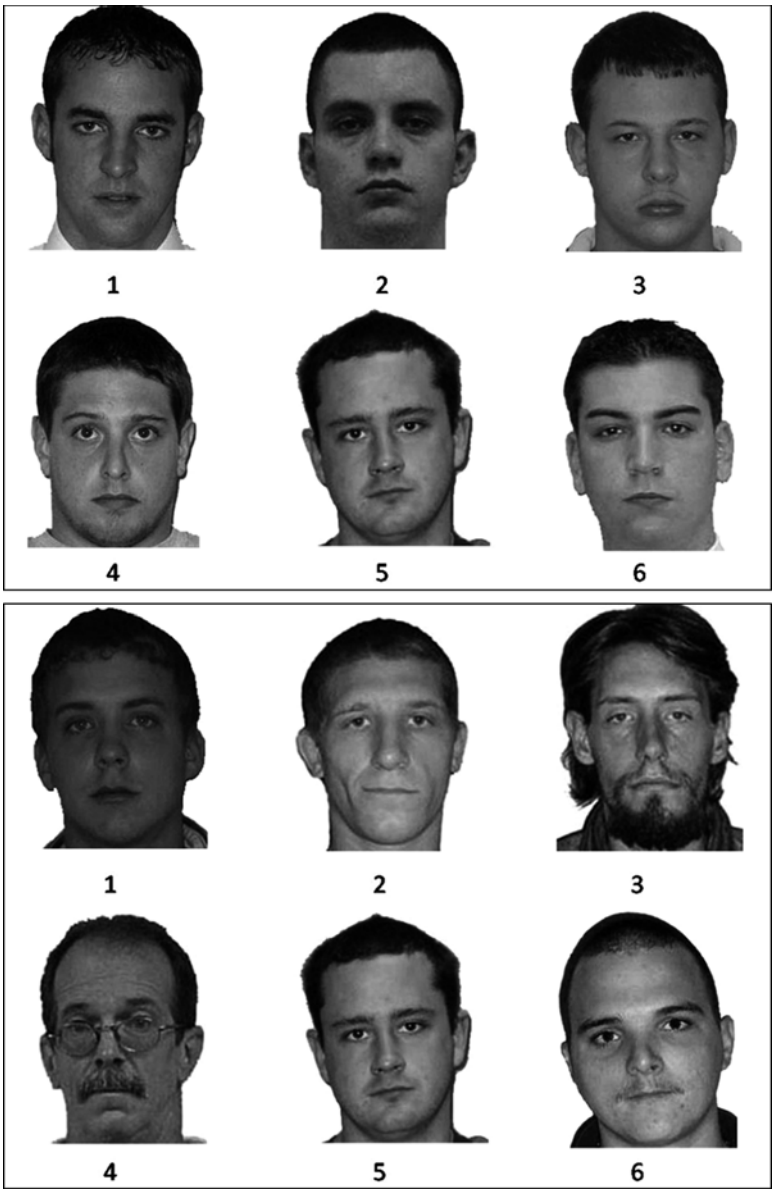


Fig. 3 Examples of a fair (*top*) and biased lineup (*bottom*). The innocent suspect is in position 5 in both lineups

the greater range over which the confidence criteria are spread (the dashed lines versus the solid vertical lines).

In summary, the data from both the meta-analysis of traditional measures and the ROC analyses paint a grim picture of showups as an identification procedure.

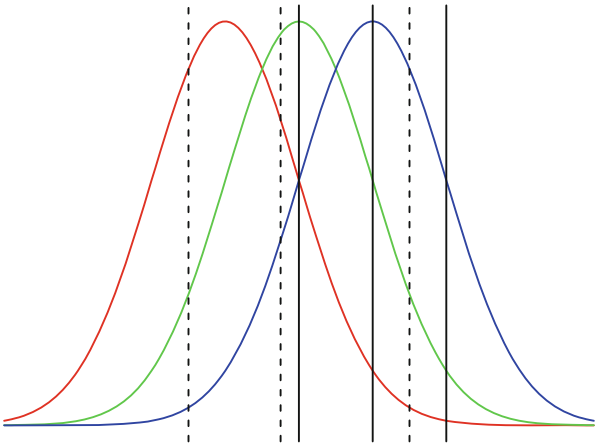


Fig. 4 Memory strength distributions depicting situations involving biased and fair lineups. The *blue* distribution represents the range of memory strengths for the perpetrator. The *green* distribution represents the varying degrees to which an innocent suspect matches participants' memory for the perpetrator. If a lineup is fair, the fillers do not stand out from an innocent suspect, akin to being selected from the green distribution. If the lineup is biased, the fillers are selected from the red distribution, which depicts a poorer match (on average) to the perpetrator. The extent of the ROC curve arises from the placement of the confidence criteria (the *vertical lines*). The greater range over which the ROC curve extends is due to the greater expanse over which the confidence criteria (the *dashed lines*) are spread if the fillers are poor (the blue distribution vs. the red distribution) than if the fillers are good matches (the blue distribution vs. the green distribution, the *solid vertical lines*)

This was true even when showups were compared to biased lineups. The only instance in which showup performance was equivalent to a lineup condition involved a sequential lineup with the suspect early in the lineup. Furthermore, even in situations in which memory theory would predict better performance in showups (i.e., immediate testing, similar cues at study and test), the performance in lineups still exceeded that of showups. Furthermore, one explanation of the superior performance in lineups to showups has been that in lineups there are filler choices, so that false identifications are spread amongst fillers rather than just to the innocent suspect, as is in the case with showups. This explanation seems faulty for two reasons. First, if there were fewer false identifications in the target absent conditions because of filler identifications, there should also be a concomitant decrease in correct identifications from target present conditions. We did not see this in our data (see Table 1), as the correct identification rate for showups was very similar to lineups (*M = .45 and M = .51, respectively*). Second, the ROC analysis clearly shows that discrimination, not bias (i.e., choosing fillers), is superior in lineups than showups (see Fig. 1, panel a).

One caveat to these conclusions is that one research team collected all of the data in the ROC analyses, and most studies utilized only one set of stimuli. Nevertheless, we believe that it is unlikely, given the large number and variety of participants, the range of independent and dependent variables used, and the supporting role provided

by theory, that the results are stimulus specific. Given the public policy implications, it is important for independent research teams to replicate the results using different materials, samples, and methodologies, before accepting the conclusions as definitive.

Confidence

The poorer performance yielded by showups might be less problematic if witnesses were not confident in their identifications. That is, perhaps the lower confidence identifications arising from a procedure deemed "suggestive," even by the U.S. Supreme Court, would not proceed to trial due to witness uncertainty. However, the data reported above signal that participants are willing to make identifications from showups with high confidence. This necessitates further investigation of how confidence relates to identification accuracy by assessing the confidence–accuracy relationship for showups and lineups.

The confidence–accuracy correlation is one of the most widely studied topics in forensic psychology (see Roediger, Wixted, & DeSoto, 2012). There are likely two major reasons for the amount of attention and research devoted to this relationship. First, under the *Biggers* criteria, jurors are instructed to use witness confidence as one index of identification accuracy (see *Neil v. Biggers*, 1972; Wells & Quinlivan, 2009). Moreover, jurors not only rely heavily on witness confidence as a proxy for accuracy, they also are more likely to excuse inconsistencies in testimony of witnesses who are highly confident in their identifications (Brewer & Burke, 2002; Krug, 2007). Second, past research has consistently shown that the correlation between confidence and accuracy for eyewitnesses is low to moderate (see Sporer, Penrod, Read, & Cutler, 1995). This is counterintuitive, as one would logically think that witnesses who display high confidence would be more accurate in their identification decisions. In fact, other research domains, such as general and text-based knowledge, show a strong positive correlation between confidence and accuracy (Stephenson, 1984; Stephenson, Clark, & Wade, 1986).

Recently, several researchers (Juslin, Olsson, & Winman, 1996; Roediger et al., 2012) have argued that the inconsistent results across research domains are likely due to the statistic used to measure the correlation. Psycho-legal researchers typically have employed the point-biserial correlation to measure the relationship between confidence and accuracy (see Roediger et al., 2012). Juslin et al. (1996) demonstrated that this measure can be misleading, as it compares confidence with the unreasonable standard of perfect discrimination. That is, all correct responses should be in one confidence category and all incorrect responses should be in the other confidence category. Furthermore, they argued that the point-biserial correlation would be low in situations where participants are presented with one stimulus and make one choice amongst similar alternatives. This is, of course, exactly the scenario used in the typical eyewitness experiment. Juslin et al. (1996) and Roediger et al. (2012) argued in favor of confidence calibration as a more

Table 2 Point-Biserial correlations for confidence and accuracy

	Non-choosers and choosers		Choosers		Suspect identifications	
Identification type	*r*	*N*	*r*	*N*	*r*	*N*
Showups	0.093^{A}	4388	0.143^{C}	1349	0.143^{E}	1349
Lineup (all)	0.152^{B}	3569	0.328^{D}	2254	0.307^{F}	1221
Simultaneous	0.144^{B}	2624	0.318^{D}	1790	0.292^{F}	947
Sequential	0.193^{B}	945	0.352^{D}	514	0.271^{F}	274
Fair lineup	0.102	1483	0.269	948	0.319^{F}	410
Bias lineup	0.162^{B}	1466	0.381^{D}	917	0.309^{F}	660
Fair simultaneous	0.93	1168	0.259	776	0.332^{F}	332
Fair sequential	0.165^{B}	315	0.310	172	0.240	78
Biased simultaneous	0.149^{B}	1148	0.352	752	0.281^{F}	524
Biased sequential	0.190	318	0.307	166	0.193	136
Overall	0.145^{B}	7877	0.291^{A}	3603	0.179^{B}	2570

Note. All identification types are compared within and overall correlations are compared between columns; values with different *superscripts* are significantly different

appropriate measure of the relationship between confidence and accuracy. Specifically, they argue that this measure allows researchers to answer the forensically relevant question: Are higher ratings of confidence associated with more accurate identifications?

With these considerations in mind, we examined the confidence–accuracy relationship for simultaneous and sequential lineups, and showups, using point-biserial correlation and calibration curves. Is the confidence–accuracy relationship weaker for showups than for lineups? To our knowledge, there have been no comparisons of this relationship across the different identification procedures.

Point-Biserial Correlation

Sporer et al. (1995) argued that it makes the most sense from a forensic standpoint to evaluate the confidence–accuracy relationship for only those people that make a choice from the identification procedure because these are the witnesses most likely to be called to testify at trial. These individuals are referred to as "choosers." In a showup, a witness can either choose the suspect or not; in lineups, witnesses also can choose fillers. The point-biserial correlations in Table 2 include the data from the same studies that contributed to the ROC analyses, as these were the only studies for which we had, or could get, confidence data. The right panel of Table 2 displays the correlations for choosers. Collapsing over identification procedure yielded a higher correlation for choosers ($r = .29$) than when choosers and non-choosers were combined ($r = .14$, $Z = 7.64$, $p < .01$). Our data mirror the results of Sporer et al. (1995), who also found that the confidence–accuracy correlation for choosers was

higher ($r = .40$) than when choosers and non-choosers were combined ($r = .29$). The confidence–accuracy correlation was significantly higher for simultaneous and sequential lineups combined than for showups, $Z = 5.71$, $p < .01$. Separately, simultaneous and sequential lineups each had significantly higher correlations than showups ($Zs > 4.31$, $ps < .01$). Thus, for choosers, the confidence–accuracy correlation is significantly greater for lineups than showups, as assessed by the point-biserial correlation.

The problem with the previous analysis is that it takes into account all choosers, including filler identifications from lineups. Filler identifications distort the calibration for lineups and puts them at a disadvantage in the same way that it complicates interpretation of identification rates (Clark & Godfrey, 2009). To address this bias, we calculated the correlations for only suspect identifications. Wixted, Mickes, Clark, Gronlund, and Roediger (2015) argued that the most forensically relevant choices are those of the perpetrator and the innocent suspect. Analyzing only suspect identifications yielded a confidence–accuracy correlation ($r = .18$) similar to that of the inclusion of all choices ($r = .14$, $Z = 1.54$, $p = .12$). Like the analysis of choosers, the confidence–accuracy correlation was significantly higher for simultaneous and sequential lineups combined than for showups, $Z = 4.38$, $p > .01$. Separately, simultaneous and sequential lineups each had higher confidence–accuracy correlations than showups ($Z = 3.69$, $p < .01$ and $Z = 2.01$, $p < .04$, respectively). Thus, for those who made suspect identifications, the confidence–accuracy correlation, albeit still low, was significantly better for lineups than showups.

Confidence Calibration

Calibration refers to how well a witness's confidence reflects his or her accuracy. Point-biserial correlations can be misleading; Juslin et al. (1996) showed that point-biserial correlations ranging from 0 to 1 could exhibit perfect calibration. The important point, as noted by Wixted et al. (2015), is that low point-biserial correlations do not necessarily imply poor confidence–accuracy relationships. The true magnitude of the relationship between confidence and accuracy is revealed by calibration curves.

Calibration curves are created by calculating the proportion of accurate decisions at each level of confidence. Perfect calibration occurs when all decisions made with 100 % confidence are accurate, decisions made with 90 % confidence are accurate 90 % of the time, and so forth. Typically, one reports several different statistics that accompany the calibration curves, which summarizes, in one form or another, how confidence relates to accuracy (Sauer, Brewer, Zweck, & Weber, 2010). The calibration index (CI) is the weighted average of the squared difference between confidence and accuracy at each confidence level. It can range from 0, representing perfect calibration, to 1, which represents no calibration. The over/under confidence (O/U) statistic measures participants' tendency to respond with greater or lesser confidence than their accuracy. For example, people who give confidence estimates of

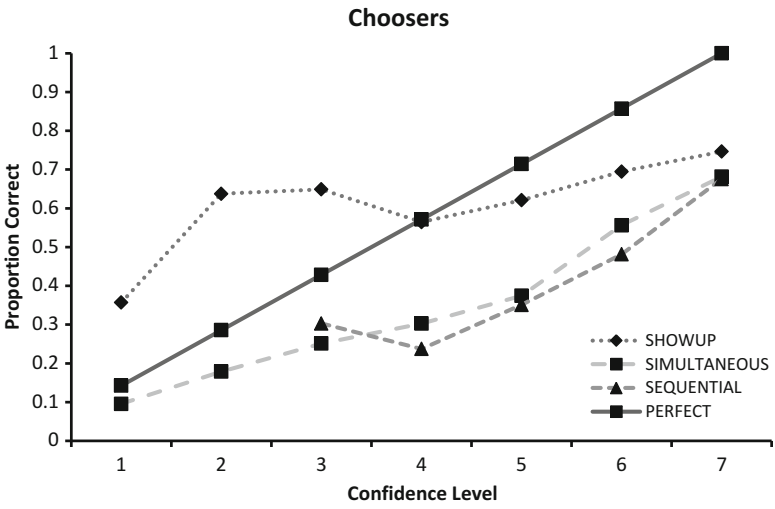


Fig. 5 Confidence calibration curves for choosers comparing simultaneous lineups, sequential lineups, and showups

30% but are only accurate 10% of the time are overestimating their confidence. O/U ranges from +1 to −1, with positive numbers representing overconfidence and negative numbers representing under confidence.

Before presenting the curves, it is important to discuss how the curves were generated. All of the lineups used a confidence scale from 1 (least confident) to 7 (most confident). The scale was converted to proportions by dividing each scale value by 7, so that the first entry is 1/7 (.14), the second is 2/7 (.29), and so on. We used the same scale for all the showup studies except Wetmore et al. (2015b), which had confidence ranging from 1 (least confident) to 10 (most confident). For this study, we combined confidence categories 2 and 3, 5 and 6, and 8 and 9 to put confidence on a 7-point scale.[7] The perfect calibration line starts at .14 to reflect the proportions on the scale. We calculated CI and O/U separately for the showups on the different scales (1–7 and 1–10), and then calculated a weighted average. All of our calibrations are between subjects because we only had participants make one identification decision/confidence judgment.

Figure 5 displays the relationship between confidence and accuracy for choosers only. Simultaneous and sequential lineup witnesses were overconfident at every level of confidence; showup witnesses, on the other hand, were underconfident at low levels of confidence and overconfident at high levels of confidence. But more to the point, note that the proportion correct from showups varies little from a confidence level of 2 (.64) to a confidence level of 7 (.75). In contrast, the proportion

[7] We reanalyzed the data excluding the Wetmore et al. (2015b) study, which was the only study using a 10-point scale. Doing this did not change the pattern of results. Consequently, we elected to leave the Wetmore et al. (2015b) in the analyses in the text.

Table 3 Calibration Index and over/under confidence for identification tasks

	Choosers		Suspect	
Identification task	CI	O/U	CI	O/U
Simultaneous	.078	−0.10	.008	−0.18
Sequential	.11	−0.06	.011	−0.12
Showup	.027	0.07	.027	0.07
Fair lineup	.03	−0.21	.017	−0.21
Biased lineup	.03	−0.14	.009	−0.14

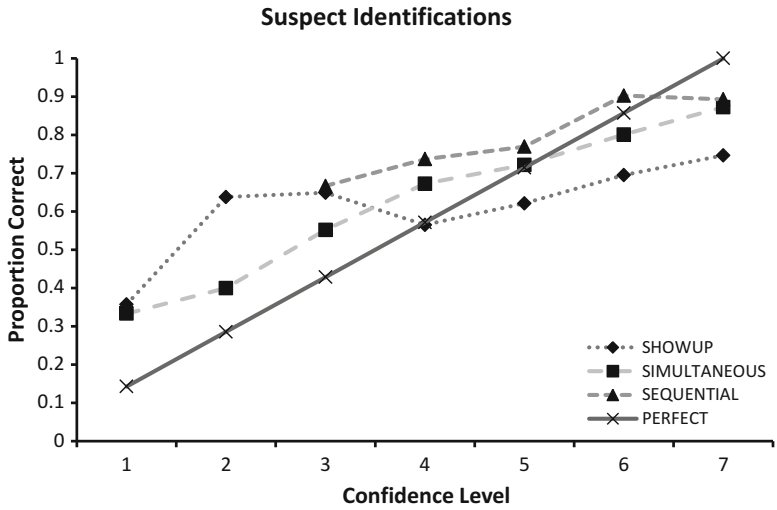


Fig. 6 Confidence calibration curves for suspect identifications comparing simultaneous lineups, sequential lineups, and showups

correct from simultaneous lineups over this same range, varies from .18 to .68. Table 3 shows that showups are better calibrated, and also indicate a penchant for overconfidence (O/U) from showups, but underconfidence from lineups. This could be problematic because the witnesses who report the highest level of confidence in their identification decisions are more likely to be called to testify in court, but from showups, these individuals were less accurate than lineup respondents.

A better comparison of showups and lineups excludes fillers and focuses on suspect identifications (see Fig. 6). Of course, the showup calibration curve does not change because all showup choosers have made a suspect identification. Table 3 now reveals the calibration of lineups to be superior to showups. We also assessed the calibration of suspect identifications from showups to fair and biased lineups (see Fig. 7). Now both biased and fair lineups were calibrated to the same degree as showups, but both still exhibited general underconfidence compared to showups.

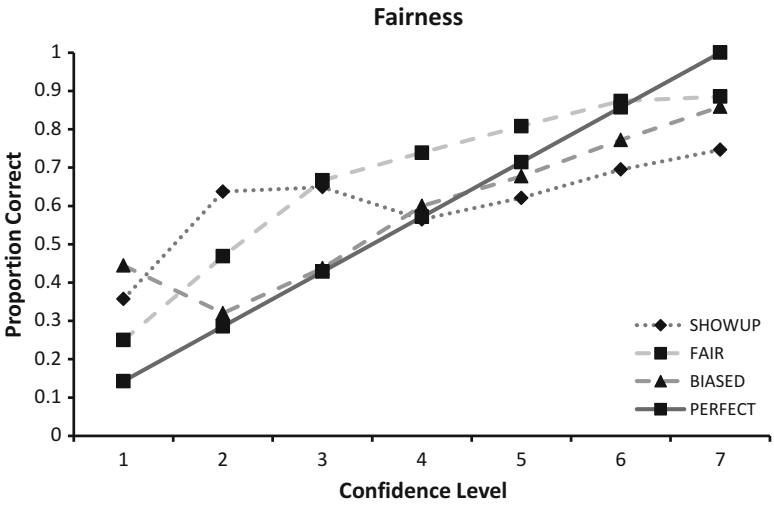


Fig. 7 Confidence calibration curves for suspect identifications comparing fair and biased lineups and showups

Conclusions

The goal of this chapter was to determine if there were situations in which a showup would outperform a lineup, thereby justifying the (occasional) use of showups. In every ROC comparison, lineups provided better discriminability and were the superior identification procedure. This was true regardless of lineup presentation method (simultaneous or sequential) and lineup fairness (fair or biased). Even in those situations for which showups might be thought to have a memorial advantage—at short retention intervals and when encoding and retrieval match—showups were worse. In sum, we have yet to find a situation where it would be more appropriate to conduct a showup if eyewitness accuracy is the primary goal. The current findings presented here provide a dismal portrayal of the most commonly used identification task.

The confidence data provide no solace for showups. When point-biserial correlations were calculated, the confidence–accuracy relationships were lower for showups than for lineups. More importantly, the forensically relevant suspect identification curves were better calibrated for lineups than showups. Witnesses viewing showups underestimated confidence at low levels and overestimated confidence at high levels. As mentioned earlier, confident eyewitnesses are most likely to testify in court: That means that for showups, the high confidence suspect identifications reflect eyewitness accuracy less well.

Wixted and Mickes (2014) provide a theoretical rationale for why showup performance is worse than lineup performance. They proposed a signal-detection-based diagnostic-feature model, arguing that simultaneous lineup performance is

superior because witnesses can compare characteristics amongst the faces. This comparison process allows eyewitnesses to identify which characteristics are shared by all suspects (and thus not helpful for distinguishing the perpetrator), and those that are specific to the perpetrator. For example, a witness may look at a simultaneous lineup and determine that all lineup members are Caucasian with short brown hair; this redundancy makes race and hair color irrelevant for identifying the perpetrator. However, shifting attention away from these irrelevant (or redundant) characteristics allows an eyewitness to focus on more diagnostic characteristics (like face shape or the width of the eyes) that likely are unique to the perpetrator. Because only one face is viewed in a showup, this comparison process is not possible; eyewitnesses are unable to determine which characteristics are irrelevant and should be ignored, and which characteristics are diagnostic and worthy of attention. Poorer discriminability arises from utilizing the wrong cues. Consequently, according to Wixted and Mickes, performance from simultaneous lineups should exceed that of showups.

This is not to say that showups are the same as one-person sequential lineups; in fact, there are several sources of evidence that argue against this interpretation. Gonzalez et al. (1993) compared showup performance to a biased lineup with only one viable option and found that a showup was not equivalent. When we have compared showups to biased sequential lineups (Gronlund et al., 2009) and simultaneous lineups, biased lineups consistently outperformed showups. Witnesses exhibit better discrimination from a biased lineup with only one viable option than from a showup. According to the Wixted and Mickes (2014) hypothesis, this must mean that witnesses have some ability to determine what features to focus on even if viewing a lineup whose members do not resemble the perpetrator.

The Wixted and Mickes (2014) theory also predicts that an eyewitness should be better able to determine which characteristics are irrelevant, and which are diagnostic, from a fair than from a biased lineup. Because the fillers in a fair lineup, by definition, share more characteristics with the perpetrator, an eyewitness should be better able to discern irrelevant characteristics and focus on the key characteristics that point to the perpetrator. For example, in a fair lineup, all lineup members are about 30, Hispanic, with short dark hair, and skinny, allowing a witness to focus in on the narrow nose of the perpetrator. However, in a biased lineup, three lineup members are the right age, but three are not. A different three have short dark hair, two have shaved heads, and one has long hair. Two are heavyset, two are moderate in weight, and two are skinny. In this situation, it is more challenging to find the right combination of cues on which to focus. Although the present data show no discriminability difference as a function of lineup fairness, contrary to this theory, additional empirical work should be undertaken to verify this surprising result.

Another prediction that arises from the Wixted and Mickes (2014) theory is that an eyewitness in a sequential lineup can determine what characteristics are diagnostic after viewing several lineup members (see also Goodsell, Gronlund, & Carlson, 2010). For example, while viewing lineup member #1, a witness notes that the ears match the perpetrator but the nose is too big. After viewing a second lineup member, a witness gets a better sense of what the nose looked like, and perhaps remembers

the perpetrator's heavy eyebrows, and so on. Consequently, discriminability improves as an eyewitness progresses through a sequential lineup. Indeed, this is exactly what Carlson et al. (2008) and Gronlund et al. (2009) found. Using the WITNESS computational memory model (Clark, 2003), Goodsell, Gronlund, and Buttaccio (2010) implemented several explanations for why a sequential lineup performance advantage seemed to arise when a suspect (guilty or innocent) was placed later in a sequential lineup. Specifically, Goodsell et al. modified WITNESS to fit data from two studies (Lindsay, Lea, & Fulford, 1991; Lindsay et al., 1991) that showed large sequential lineup advantages, both of which involved the placement of suspects into position 8 of 8. One of Goodsell et al.'s explanations in particular, the better cue model, was very similar to the ideas expressed in Wixted and Mickes.

There are many benefits to providing theoretical explanations for empirical findings (e.g., Bjork, 1973; Brewer, Weber, & Semmler, 2007; Clark, 2008). Including requiring researchers to operationalize and specify constructs of interest, viewing data in a more rigorous manner, and encouraging the generation of testable hypotheses (see Farrell & Lewandowsky, 2010; see also Clark & Gronlund, 2015). Only by evaluating data in the context of well-specified, testable theories can we make cumulative progress in understanding how eyewitnesses make decisions and what identification procedures might best support that decision-making.

One overlooked benefit of providing theoretical explanations is that a theoretical idea might point to a new method by which we can enhance performance. For example, if participants indeed can learn as a sequential lineup progresses, previewing a series of known innocent faces prior to a showup might engender the same comparative processes that allow an eyewitness to focus on diagnostic cues and ignore irrelevant cues. Goodsell, Gronlund and Buttaccio (2010) had participants view a mock crime and then evaluate known innocent faces just prior to a lineup identification. This pre-ID procedure showed beneficial effects for witnesses that evaluated faces with a high degree of match to the perpetrator prior to making an identification from a simultaneous lineup. It is important to determine if a method like this can enhance showup discriminability, as it seems likely that, despite the overwhelming evidence favoring lineups over showups, law enforcement will continue to use showups in the field.

In conclusion, the goal of this chapter was to compare showups and lineups using a variety of measures and lineup procedures (e.g., simultaneous, sequential, fair, biased). In all of our comparisons, we found that showups were at a disadvantage relative to lineups. More specifically, showups and lineups yielded similar correct identifications but showups consistently had more false identifications, as represented by significantly lower ROC curves. Even in situations in which showups should have, based on memory theory, a memorial advantage to lineups (e.g., clothing match and shorter retention interval), they still underperformed lineups. Furthermore, the calibration curves indicated that not only were showup witnesses less accurate, but at the highest levels of confidence, they were overconfident in the accuracy of their identification choices. It is likely that in forensic situations, those witnesses who have the highest confidence will be the ones that testify at trial. In our data, these were the most dangerous showup witnesses because although they were

inaccurate they were highly confident. To the extent that the practice of showups is going to continue, and we believe it will, it is important to investigate ways to make this procedure yield more accurate identification choices.

Author Notes Jeffrey S. Neuschatz, Department of Psychology, The University of Alabama in Huntsville; Scott D. Gronlund and Stacy Wetmore, Department of Psychology, University of Oklahoma; Charles A. Goodsell, Department of Psychology, Canisius College. This work was supported by the National Science Foundation grant SES-1060921 to JSN, SES-1060902 to SDG, and SES-1060913 to CAG. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of the NSF.

References

Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior, 25(5)*, 475–491. doi:10.1023/A:1012840831846.

Bjork, R. A. (1973). Why mathematical models? *American Psychologist, 28(5)*, 426–433. doi:10.1037/h0034623.

Brewer, N., & Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and Human Behavior, 26*(3), 353–364. doi:10.1023/A:1015380522722.

Brewer, N., Weber, N., & Semmler, C. (2007). A role for theory in eyewitness identification research. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology* (Memory for people, Vol. II, pp. 201–218). Mahwah, NJ: Lawrence Erlbaum.

Carlson, C. A., & Carlson, M. A. (2014). An evaluation of lineup presentation, weapon presence, and a distinctive feature using ROC analysis. *Journal of Applied Research in Memory and Cognition, 3*, 45–53. doi:10.1016/j.jarmac.2014.03.004.

Carlson, C. A., Gronlund, S. D., & Clark, S. E. (2008). Lineup composition, suspect position, and the sequential lineup advantage. *Journal of Experimental Psychology: Applied, 14*, 118–128. doi:10.1037/1076-898X.14.2.118.

Clark, S. E. (2003). A memory and decision model for eyewitness identification. *Applied Cognitive Psychology, 17*, 629–654.

Clark, S. E. (2008). The importance (necessity) of computational modeling for eyewitness identification research. *Applied Cognitive Psychology, 22*, 803–813. doi:10.1002/acp.1484.

Clark, S. E. (2012). Costs and benefits of eyewitness identification reform: Psychological science and public policy. *Perspectives on Psychological Science, 7*(3), 238–259. doi:10.1177/174569161239584.

Clark, S. E., & Godfrey, R. D. (2009). Eyewitness identification evidence and innocence risk. *Psychonomic Bulletin & Review, 16*, 22–42. doi:10.3758/PBR.16.1.22.

Clark, S. E., & Gronlund, S. D. (2015). Mathematical modeling shows that compelling stories do not make for accurate descriptions of data. In J. G. W. Raaijmakers, R. Goldstone, M. Steyvers, A. Criss, & R. M. Nosofsky (Eds.), *Cognitive modeling in perception and memory: A festschrift for Richard M. Shiffrin*. New York, NY: Psychology Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Dekle, D. J. (1997). Testing delays resulting in increased identification accuracy in line-ups and show-ups. *Journal of Offender Rehabilitation, 25*, 35–49. doi:10.1300/J076v25n03_03.

Dekle, D. J., Beal, C. R., Elliott, R., & Huneycutt, D. (1996). Children as witnesses: A comparison of lineup versus showup identification methods. *Applied Cognitive Psychology, 10*(1), 1–12. doi:10.1002/(SICI)1099-0720(199602)10:1<1::AID-ACP354>3.0.CO;2-Y.

Dobolyi, D. G., & Dodson, C. S. (2013). Eyewitness confidence in simultaneous and sequential lineups: A criterion shift account for sequential mistaken identification overconfidence. *Journal of Experimental Psychology: Applied, 19*(4), 345–357. doi:10.1037/a0034596.

Dysart, J. E., & Lindsay, R. C. L. (2007). Show-up identifications: Suggestive technique or reliable method? In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology* (Memory for people, Vol. II, pp. 137–153). Mahwah, NJ: Lawrence Erlbaum.

Dysart, J. E., Lindsay, R. C. L., & Dupuis, P. R. (2006). Show-ups: The critical issue of clothing bias. *Applied Cognitive Psychology, 20*, 1009–1023. doi:10.1002/acp.1241.

Farrell, S., & Lewandowsky, S. (2010). Computational models as aids to better reasoning in psychology. *Current Directions in Psychological Science, 19*(5), 329–335.

Gardner, T. J., & Anderson, T. M. (2004). *Criminal evidence: Principle and cases* (5th ed.). Belmont, CA: Thompson Wadsworth.

Garrett, B. (2011). *Convicting the innocent: Where criminal prosecutors go wrong*. Cambridge, MA: Harvard University Press.

Garrioch, L., & Brimacombe, E. C. A. (2001). Lineup administrators' expectations: Their impact on eyewitness confidence. *Law and Human Behavior, 25*, 299–315. doi:10.1023/A:1010750028643.

Gonzalez, R., Ellsworth, P. C., & Pembroke, M. (1993). Response biases in lineups and showups. *Journal of Personality and Social Psychology, 64*, 525–537. doi:10.1037/0022-3514.64.4.525.

Goodsell, C. A., Gronlund, S. D., & Buttaccio, D. (2010). *A test of the better memory probe hypothesis: Improving eyewitness identification accuracy*. Presented at the AP-LS meetings, Vancouver, March, 2010.

Goodsell, C. A., Gronlund, S. D., & Carlson, C. A. (2010). Exploring the sequential lineup advantage using WITNESS. *Law and Human Behavior, 34*, 445–459. doi:10.1007/s10979-009-9215-7.

Goodsell, C. A., Wetmore, S. A., Neuschatz, J. S., & Gronlund, S. D. (2013). Showups. In B. L. Cutler (Ed.), *Reform of eyewitness identification procedures* (pp. 45–63). Washington, DC: American Psychological Association.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford, England: Wiley.

Gronlund, S. D., Andersen, S. M., & Perry, C. (2013). Presentation methods. In B. Cutler (Ed.), *Reform of eyewitness identification procedures*. Washington, DC: APA Publications.

Gronlund, S. D., Carlson, C. A., Dailey, S. B., & Goodsell, C. A. (2009). Robustness of the sequential lineup advantage. *Journal of Experimental Psychology: Applied, 15*, 140–152. doi:10.1037/a0015082.

Gronlund, S. D., Carlson, C. A., Neuschatz, J. S., Goodsell, C. A., Wetmore, S. A., Wooten, A., & Graham, M. (2012). Showups versus lineups: An evaluation using ROC analysis. *Journal of Applied Research in Memory and Cognition*, *1*, 221–228. doi: 10.1016/j.jarmac.2012.09.003.

Gronlund, S. D., & Neuschatz, J. S. (2014). Eyewitness identification discriminability: ROC analysis versus logistic regression. *Journal of Applied Research in Memory and Cognition, 4*, 221–228.

Gronlund, S. D., Wixted, J. T., & Mickes, L. (2014). Evaluating eyewitness identification procedures using ROC analysis. *Current Directions in Psychological Science, 23*, 3–10. doi:10.1177/0963721413498891.

Haw, R. M., & Fisher, R. P. (2004). Effects of administrator-witness contact on eyewitness identification accuracy. *Journal of Applied Psychology, 89*, 1106–1112. doi:10.1037/0021-9010.89.6.1106.

Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence–accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1304. doi:10.1037/0278-7393.22.5.1304.

Kassin, S. M., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the "general acceptance" of eyewitness testimony research: A new survey of the experts. *American Psychologist, 56*, 405–416. doi:10.1037/0003-066X.56.5.405.

Key, K. N., Cash, D. K., Neuschatz, J. S., Price, J. L., Wetmore, S. A., & Gronlund, S. D. (2015). Age differences (or lack thereof) in discriminability for lineups and showups. *Psychology, Crime, & Law, 21*(9), 871–889. doi:10.1080/1068316X.2015.1054387.

Krug, K. (2007). The relationship between confidence and accuracy: Current thoughts of the literature and a new area of research. *Applied Psychology in Criminal Justice, 3*, 7–41.

Lawson, V. Z., & Dysart, J. E. (2014). The showup identification procedure: An exploration of systematic biases. *Legal and Criminalogical Psychology, 19,* 54–68. doi:10.1111/j.2044-8333.2012.02057.x.

Light, L. L. (1996). Memory and aging. In E. L. Bjork & R. A. Bjork (Eds.), *Memory: Handbook of perception and cognition* (2nd ed., pp. 444–477). San Diego, CA: Academic.

Lindsay, R. C. L., Lea, J. A., & Fulford, J. A. (1991). Sequential lineup presentation: Technique matters. *Journal of Applied Psychology, 76*, 741–745. doi:10.1037/0021-9010.76.5.741.

Lindsay, R. C. L., Pozzulo, J. D., Craig, W., Lee, K., & Corber, S. (1997). Simultaneous lineups, sequential lineups, and showups: Eyewitness identification decisions of adults and children. *Law and Human Behavior, 21*, 391–404. doi:10.1023/A:1024807202926.

Luus, E. C. A., & Wells, G. L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behavior, 15*, 43–57. doi:10.1007/BF01044829.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Malpass, R. S., & Lindsay, R. C. L. (1999). Measuring lineup fairness. *Applied Cognitive Psychology, 13*, 1–8.

McQuiston, D., & Malpass, R. (2001, June). *Eyewitness identifications in criminal cases: An archival study*. Paper presented at the fourth biennial meeting of the Society for Applied Research in Memory and Cognition, Kingston, ON.

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory & Cognition, 33*, 783–792. doi:10.3758/BF03193074.

Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: Comparing diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied, 18*, 361–376. doi:10.1037/a0030609.

Neil v. Biggers. (1972). 409 U.S. 188.

Quinlivan, D. S., Neuschatz, J. S., Cutler, B. L., Wells, G. L., McClung, J., & Harker, D. L. (2011). Do pre-admonition suggestions moderate the effect of unbiased lineup instructions? *Legal and Criminological Psychology, 17*, 165–176. doi:10.1348/135532510X533554.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics, 12*(1), 77. http://dx.doi.org/10.1186/1471-2105-12-77.

Roediger, H. L., Wixted, J. H., & DeSoto, K. A. (2012). The curious complexity between confidence and accuracy in reports from memory. In L. Nadel & W. Sinnott-Armstrong (Eds.), *Memory and law* (pp. 84–118). New York, NY: Oxford University Press.

Rosenthal, R. (1966). *Experimenter effects in behavioral research*. New York, NY: Appleton.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York, NY: Holt Reinhart and Winston.

Sauer, J., Brewer, N., Zweck, T., & Weber, N. (2010). The effect of retention interval on the confidence–accuracy relationship for eyewitness identification. *Law and Human Behavior, 34*, 337–347. doi:10.1007/s10979-009-9192-x.

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin, 118(3)*, 315–327. doi:10.1037/0033-2909.118.3.315.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2003). Eyewitness accuracy rates in police showup and lineup presentations: A meta-analytic comparison. *Law and Human Behavior, 27*, 523–540. doi:10.1023/A:1025438223608.

Steblay, N. K., Dysart, J. E., & Wells, G. L. (2011). Seventy-two tests of the sequential lineup superiority effect: A meta-analysis and policy discussion. *Psychology, Public Policy, and Law, 17*(1), 99–139. doi:10.1037/a0021650.

Stephenson, G. M. (1984). Accuracy and confidence in testimony: A critical review and some fresh evidence. In D. J. Muller, D. E. Blackman, & A. J. Chapman (Eds.), *Psychology and law: Topics from an international conference* (pp. 229–248). Chichester, England: Wiley.

Stephenson, G. M., Clark, N. K., & Wade, G. S. (1986). Meetings make evidence? An experimental study of collaborative and individual recall of a simulated police interrogation. *Journal of Personality and Social Psychology, 50*, 1113–1122. doi:10.1037/0022-3514.50.6.1113.

Stovall v. Denno, 388 U.S. 293 (1967).

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest, 1*(1), 1–26.

Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement*. Washington, DC: United States Department of Justice, Office of Justice Programs.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373. doi:10.1037/h0020071.

United States v. Wade, 388 U. S. 293 (1967)

Valentine, T., Davis, J. P., Memon, A., & Roberts, A. (2012). Live showups and their influence on a subsequent video line-up. *Applied Cognitive Psychology, 26*, 1–23. doi:10.1002/acp.1796.

Wells, G. L., & Quinlivan, D. S. (2009). Suggestive eyewitness identification procedures and the Supreme Court's reliability test in light of eyewitness science: 30 Years later. *Law and Human Behavior, 33*, 1–24. doi:10.1007/s10979-008-9130-3.

Wetmore, S. A., Neuschatz, J. S., Gronlund, S. D, Goodsell, C. A., Wooten, A., & Carlson, C. A. (2015a). Effect of retention interval on showup and lineup performance. *Journal of Applied Research in Memory and Cognition, 4,* 8–14. doi:10.1016/j.jarmac.2014.07.003.

Wetmore, S. A., Neuschatz, J. S., Gronlund, S. D., Key, K. N., & Goodsell, C. A. (2015b). Do the clothes make the criminal in showups? *Journal of Applied Research in Memory and Cognition, 4,* 36–42. doi:10.1016/j.jarmac.2014.12.003.

Wixted, J. T., & Mickes, L. (2012). The field of eyewitness memory should abandon probative value and embrace receiver operating characteristic analysis. *Perspectives on Psychological Science, 7*, 275–278. doi:10.1177/1745691612442906.

Wixted, J. T., & Mickes, L. (2014). A signal-detection based diagnostic-feature model of eyewitness identification. *Psychological Review, 121*, 262–276. doi:10.1037/a0035940.

Wixted, J. T., Mickes, L., Clark, S. E., Gronlund, S. D., & Roediger, H. L. (2015). Initial eyewitness confidence reliably predicts eyewitness identification accuracy. *American Psychologist, 70,* 515–526. doi:10.1037/a0039510.

Yarmey, A. D., Yarmey, A. L., & Yarmey, M. J. (1994). Face and voice identifications in showups and lineups. *Applied Cognitive Psychology, 8*, 453–464. doi:10.1002/acp.2350080504.

Yarmey, A. D., Yarmey, M. J., & Yarmey, A. L. (1996). Accuracy of eyewitness identification in showups and lineups. *Law and Human Behavior, 20*, 459–477. doi:10.1007/BF014989.

RECEIVED by MSC 4/12/2022 1:07:49 PM

APPENDIX G

Do MI Sentencing Guidelines Meet the Legislature Goals

RECEIVED by MSC 4/12/2022 1:07:49 PM

Do Michigan's Sentencing Guidelines Meet the Legislature's Goals?

A Historical and Empirical Analysis of Prison Terms for Life-Maximum Offenses

Barbara Levine, Anne Mahar, & Justin Smith

November 2021

Safe & Just Michigan

www.safeandjustmichigan.org

RECEIVED by MSC 4/12/2022 1:07:49 PM

About the Authors

Barbara Levine, J.D. began her legal career representing indigent defendants on appeal at the State Appellate Defender Office before Michigan had any sentencing guidelines. She was the first administrator of the Michigan Appellate Assigned Counsel System, the executive director of Citizens Alliance on Prisons and Public Spending (the predecessor to Safe and Just Michigan), and a member of the Criminal Justice Policy Commission. She received both her Bachelor's and Juris Doctor degrees from the University of Michigan.

Anne Mahar, Ph.D. is the research specialist at Safe & Just Michigan. She previously was an Assistant Professor in the Department of Sociology, Anthropology and Criminal Justice at Arcadia University. She received her B.S. and M.A. in Criminal Justice and Criminology from Eastern Michigan University, and Ph.D. from Old Dominion University.

Justin Smith, Ph.D. is an associate professor in the Department of Sociology and Criminology at the University of North Carolina-Wilmington. He previously held the same position at Central Michigan University. His research projects include studying racial inequality, particularly as it relates to prisons and courts, the war on drugs, and the lives of the formerly incarcerated. He is an instructor for the Inside-Out Prison Exchange Program. He received his B.A. and M.A. from Western Michigan University and Ph.D. from the University of Florida.

Acknowledgements

The authors would like to thank several people whose contributions have enriched this report. This work would not have been possible without the efforts of Elsie Kettunen, who extracted relevant information from the MDOC's massive database and organized it into the datasets that underlie this report. Michigan Attorney Patricia Streeter's persistence and attention to detail made the comparison of Michigan's guidelines to those of other jurisdictions possible. Judge Douglas Shapiro, Jacqueline McCann, and Kelly Mitchell each read selected portions, and provided helpful insights. The meticulous editing of Annalisa Zox-Weaver and Jacqueline Carsello significantly improved the final product.

Note: Safe & Just Michigan avoids use of the term "offender" whenever possible. It is used in this report when adhering to legal terminology.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table of Contents

The Point of This Report 1

Section I. History and Context

Chapter One. The Evolution of Sentencing Guidelines 6

A. The Michigan Felony Sentencing Project Creates a Vision 6
B. The Judiciary Adopts Voluntary Guidelines 7
C. *Milbourn* Sets a "Proportionality" Standard for Appellate Review 10
D. The Legislature Enacts Mandatory Guidelines 11
1. Structural Differences 11
2. Raising the Minimums 13
3. Habitual Offenders 14
E. The Guidelines Are Not Assessed; Sentences Get Longer 15
1. Guidelines Amendments 15
2. Consecutive Sentences 16
3. Mandatory Minimums 18
4. A Lost Opportunity 19

Chapter Two. Changing Views of How Long is Long Enough 20

A. 25-Year Trends and Mass Incarceration 20
1. The Michigan Evidence 20
2. Criminologists Weigh In 22
B. Consequences of Longer Sentences 23
1. The Fiscal Costs 23
2. The Impact on Public Safety 24

Chapter Three. Prior Research on Michigan Sentencing Guidelines 26

A. Michigan Felony Sentencing Project (1979) 26
B. Abel F. Ekpunobi, "Judicial Decision Making Under Michigan Sentencing Guidelines" (1999) 27
C. National Center for State Courts (2008) 28
1. Background 28
2. The Impact of Guidelines Elements on Sentence Length 29
3. The Impact of Extra-Guidelines Factors 29
4. NCSC Conclusions and Recommendations 30
D. The Council of State Governments (2014) 31
E. Criminal Justice Policy Commission (2019) 31

Section II. The Current Research

Chapter Four. A Summary of the Findings 34

Chapter Five. How Sentences Grew Longer and More Numerous 48

A. Life vs. Long Indeterminate Sentences 48
B. Habitual vs. Non-Habitual Sentences 50

RECEIVED by MSC 4/12/2022 1:07:49 PM

C. Differences in Impact among Counties 52
1. Non-Habitualized Defendants 53
2. Habitualized Defendants 54
D. The Increase in Multiple Sentences Per Defendant 55

Chapter Six. How the Guidelines Promote Longer Sentences 60

A. The Grids 60
1. Sentence Increases from Grid to Grid 60
2. The Number and Redundancy of Cells 61
B. Prior Record Variables 62
C. Offense Variables 66
D. The Intersection of the PRVs and OVs on the M2 and A Grids 72
E. Broader Cell Ranges for Habitual Offenders 73
F. Bringing It All Together: Regression Analyses 75

Chapter Seven. How the Guidelines Promote Disparity 77
A. How Sentences are Distributed Within, Above, and Below the Applicable Range .. 77
1. Non-Habitual Sentences 78
2. Habitual Sentences 81
3. Departures in the Regression Analyses 83
B. The Impact of PRVs and OVs on Where Within Ranges Sentences Fall 83
1. PRV Levels 84
2. OV Levels 84

Chapter Eight. How the Actual Minimum Sentences Relate to the Recommended Ranges 86
A. The Number of Different Medians 90
B. Medians below the Minimum-Minimum 90
C. Medians at the Minimum-Minimum 91
D. Same Ranges, Different Medians: Outcomes in Redundant Cells 92
E. Different Ranges, Identical Medians: Outcomes in Non-Redundant Cells 92
F. Same Ranges, Different Offenses: Medians All Over the Map 92

SUMMARY OF OUTCOMES RESULTING FROM THE GUIDELINES STRUCTURE 94

Chapter Nine. Systemic Factors: Method of Conviction 95

A. The Risks of Plea Bargaining and the Decision to Go to Trial 95
B. Frequency of Conviction Methods 96
C. Relationship to Sentence Length: The Size of the Trial Tax 98
D. How the Trial Tax is Applied under Sentencing Guidelines 101

Chapter Ten. Systemic Factors: County of Conviction 103

A. Median Sentence 104
B. Place on the Range 105
C. Conviction Method 109
D. County-Specific Analyses 110
E. Further Observations 112

Chapter Eleven. Systemic Factors: Sentencing Judge 113

RECEIVED by MSC 4/12/2022 1:07:49 PM

Chapter Twelve. Defendant Characteristics: Sex and Age 116
A. Sex 116
B. Age 118

Chapter Thirteen. Defendant Characteristics: Race 120
A. County Population 121
B. Age 121
C. Median Sentence Length 123
D. Prior Record Level 125
E. Offense Severity Level 127
F. Conviction Method 128
G. Place on the Range 129
H. Conclusion 130

Chapter Fourteen. Habitual Offender Enhancements: Who Gets Chosen? 131
A. How Many Sentences are Eligible for the Habitual Offender Enhancement and How Many Receive It? 131
B. How are Sentences Selected for the Habitual Enhancement? 133
1. Guidelines Factors 133
2. Conviction Method 133
3. Defendant Characteristics 134
C. How Much Longer are Habitual Sentences? 134

Section III. Policy Implications

Chapter Fifteen. Comparing Michigan to Other Jurisdictions 138
A. Introduction 138
B. Overview of Six Jurisdictions 139
1. Kansas 139
2. Minnesota 140
3. North Carolina 142
4. Oregon 143
5. Utah 145
6. United States 146
C. Hypothetical Case Results 150
D. How Does Michigan's Guidelines System Compare? 155
1. Commission Oversight 155
2. Enforceability 156
3. Establishing Offense Seriousness: Grids and Offense Severity Levels 156
4. Defining Prior Conviction 157
5. Aggravation, Enhancement, and Habitualization 158
6. "Real Offense" Sentencing 159
7. Double-Counting 160

Chapter Sixteen. Should the Guidelines be Enforced? The Results of Reasonableness 162
A. How Guidelines Can be Made Mandatory 162
B. Why the Distinction Matters 163
1. Review Under the Advisory Judicial Guidelines 164
2. Review Under the Mandatory Legislative Guidelines 164
3. Review Under the Advisory Legislative Guidelines 165
4. Application of the Reasonableness Standard after *Dixon-Bey* 166

RECEIVED by MSC 4/12/2022 1:07:49 PM

C. Are Mandatory Guidelines Desirable? 169

Chapter Seventeen. Recommendations for Change 171

Conclusion 174

Endnotes 177

Appendices

Technical Appendix A. The Data and Methods of Analysis 196
A. Data 196
B. Analytic Strategies 198
1. Univariate Analysis: Examining Sentence Lengths 198
2. Bivariate Analysis: The Relationship Between Sentence Lengths and Individual Variables 200
3. Multivariate Analysis: How Multiple Factors Interact to Produce a Sentence 201

Technical Appendix B. Determining Sentence Length: Which Factors Count the Most? 203

Appendix C. Additional Tables 215
Table C1. Racial Demographics of Counties and Sentences 215
Table C2. High and Low Minimum Sentences by County and Offense 215

Appendix D. Hypothetical Case Comparison Charts 217

RECEIVED by MSC 4/12/2022 1:07:49 PM

Do Michigan's Sentencing Guidelines Meet the Legislature's Goals?

THE POINT OF THIS REPORT

Protecting citizens from crime—or at least punishing those who commit it—is one of the primary functions of state government. But views on how this responsibility should be fulfilled shift over time. Public sentiment about punishment for crime has softened since the "get tough" years of the 1980s and 1990s; public concern about the social and fiscal costs of mass incarceration has increased. Harsh prison sentences for drug crimes have been revised; alternatives to incarceration for non-violent offenses are more readily available; drug and alcohol treatment courts have become common; the use of jails and prisons as the mental health system of last resort is being decried; community support for people returning from prison has expanded. At the end of 2019, the Michigan prisoner population was at 38,053, the lowest since 1991.

However, one particularly stubborn problem persists: How to sentence individuals who commit serious crimes against other people, like murder, armed robbery, and sexual assault. These cases provoke strong emotions, involve vulnerable victims and/or family members, and often receive attention from the press. They also raise difficult questions of law and policy.

- While some amount of prison time for such offenses may be inevitable, how much is enough?
- Is there a point at which very long sentences are disproportionate to the crime, unfair to the individual, unnecessary to protect the public or an ineffective use of resources?
- How much discretion should local judges have to select the sentence? What factors should they be able to consider?
- What role should the prosecutor play in determining the sentence?
- Who should ultimately decide what sentence is appropriate and by what standard?
- How important is it that sentences for serious assaultive offenses be consistent, so that people do not receive widely varying prison terms for essentially similar conduct?

Addressing these questions is often the work of a sentencing commission. The answers may change over time as data become available and public attitudes evolve. In the absence of a commission in Michigan, this report aims to stimulate data-driven discussion. To begin, it helps to understand a few basic points about the structure of Michigan sentencing.

Striking a balance between too much discretion for decision-makers and too little is at the core of criminal justice policy-making.

Although the criminal justice process is governed by many complex rules, the implementation of those rules depends on the exercise of discretion by thousands of individuals—from police officers and prosecutors to judges and parole board members—each of whom has their own vision of the "right" sentence. Different states, different counties within states, and different judges within the same counties weight the goals of criminal sentencing differently and may reach strikingly disparate outcomes based on similar sets of facts. The entire system rests on a series of judgment calls made by multiple (often elected) decision-makers in the context of ever-evolving societal values. The tension between the need for rules to constrain the exercise of power by individual officials and the need for discretion so that these individuals can apply rules fairly and sensibly to individual defendants persists.

That tension is exacerbated by questions of local control. Michigan judges and prosecutors are locally elected; trial-level courts and jails are, for the most part, locally funded. Some argue that sentences should reflect the values of the local community and the predilections of local officials. Yet defendants are being convicted of violating state law and sentenced by judges whose salaries are paid by the state. And the prisons that judges use to carry out their sentences are operated at the expense of state taxpayers. Ensuring reasonable consistency in how the law is applied, how punishment is apportioned, and how resources are expended is an ideal but difficult goal.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Sentencing guidelines are meant to help strike that balance in accordance with stated policy goals.

Many states utilize sentencing guidelines that are meant to help judges choose sentences that are consistent, unbiased, proportional to both the offense and the defendant, and realistic in their use of scarce resources. Michigan's guidelines establish ranges within which minimum sentences should fall depending on the defendant's criminal history and the seriousness of the offense.

Three points about the legislative guidelines enacted in 1998 indicate that now is the time to reconsider them.

1. They were developed during the period in which numerous "tough on crime" measures were being adopted. In fact, their passage was tie-barred to the elimination of all good behavior credits for prisoners in the name of "truth in sentencing." Part of their stated purpose was to increase prison terms for the most serious crimes.

2. Little systematic review of how well the guidelines are meeting their goals has been conducted. Unlike other jurisdictions with active, well-staffed commissions, the original sentencing commission was disbanded shortly after the current guidelines were adopted. A subsequent commission also came and went without conducting a comprehensive analysis.[1] In the meantime, the Legislature has repeatedly amended the guidelines on an ad hoc basis.

3. As enacted by the Legislature compliance with the guidelines recommendation was mandatory unless the sentencing judge stated "substantial and compelling reasons" for departing. In a 2015 decision, the Michigan Supreme Court made the guidelines advisory, and the sentence imposed must now simply be "reasonable."

Numerous choices made by legislators, prosecutors, and appellate courts expand or contract the boundaries within which judges act. Moreover, the minimum sentence a judge imposes is only one determinant of how long the defendant will actually spend in prison.[2] Nonetheless, the selection of the minimum sentence is a critical component of the criminal justice process. It not only has obvious consequences for defendants, victims, criminal justice officials, and taxpayers, but also represents the community's values regarding proportionality and consistency of punishment.

Life-maximum offenses afford sentencing judges enormous discretion and have a major impact on the prison population.

Under Michigan's penal code, most sentences are "indeterminate." The Legislature sets the maximum term, and a judge sets the minimum, which cannot exceed two-thirds of the maximum. Of the prison sentences imposed in 2012, 45% carried maximums of five years or less, and 75% carried maximums no greater than 15 years. In these cases, whatever the sentencing guidelines recommended, unless the defendant was charged as a habitual offender, the breadth of the judge's discretion in setting a *minimum* sentence was ultimately constrained by the statutory *maximum* to no more than 10 years and, more often, to three and one-third years or less.

For the most serious crimes short of first-degree murder, the situation is different. The penalty for these offenses is life or any term. If the judge chooses "life," the person will become eligible for parole in 15 years. If—as is far more common—the judge chooses an indeterminate term, they select both the minimum and the maximum, which could be 10-20, 30-50, 60-90, or one of an infinite variety of other combinations.[3] Even the two-thirds rule limiting the length of minimum sentences does not apply. In 2019, 35.0% of the total Michigan prisoner population was serving for one of four life-maximum offenses: second-degree murder (M2), assault with intent to commit murder (AWIM), first-degree criminal sexual conduct (CSC), or armed robbery (RA). Roughly 6,000 people—just more than 16% of the population—were serving either parolable life sentences or minimum terms of over 20 years. The

RECEIVED by MSC 4/12/2022 1:07:49 PM

prison beds allotted to people who are serving minimum terms of one, two, or three years turn over quickly. But the people serving very long minimums stay in their beds for many years. As new long-termers enter prison, these prisoners stack up, and their share of the population keeps growing. Thus, the life-maximum offenses present the most difficult and consequential questions about how judicial discretion should be exercised and the extent to which it should be constrained.

Life-maximum offenses are peculiarly relevant to examining the operation of the sentencing guidelines for multiple reasons.

- These are the serious offenses that the legislative guidelines policy of lengthening sentences was primarily designed to affect.
- Because these cases provide judges with the maximum amount of sentencing discretion, they provide the clearest window into how and with what consequences that discretion is exercised and the extent to which the guidelines actually constrain it.
- These are the cases in which a prosecutor's decision about whether to charge someone as a habitual offender can have by far the most dramatic effect.
- These offenses have a large and growing impact on the size and cost of the prison population.
- These are the cases in which disproportionately harsh and/or inconsistent sentences have the greatest impact on individual defendants, their families, and their communities.
- Criminal justice experts both in Michigan and nationally have come to recognize that significantly reducing prison populations will be impossible without addressing the very long sentences imposed on people convicted of assaultive and sex offenses.

This report examines how the guidelines have affected sentence lengths and whether they have achieved the goals of increased proportionality and decreased disparity in sentences for life-maximum crimes.

Section I places the Michigan guidelines in historical and practical perspective. Chapter One traces the evolution of the guidelines, describes their enforcement by the appellate courts, and then places them in the context of other legislative changes that have lengthened sentences. Chapter Two examines the extent to which sentences for life-maximum offenses have grown longer over the decades and the consequences of that trend. Chapter Three reviews the prior research on the need for and effect of Michigan's sentencing guidelines.

Section II presents an empirical analysis of more than 38,000 sentences imposed for the four most common life-maximum offenses in 13 Michigan counties over the course of nearly 30 years. The research examines:

- How sentences for the life-maximum offenses grew when the judicial guidelines were replaced by legislative guidelines.
- How the structure of the legislative guidelines, including the complex series of grids, the prior record and offense variables, the breadth of the recommended ranges, and the extent of both upward and downward departures from those ranges, permit or promote certain outcomes.
- How effective the legislative guidelines have been in reducing disparity based on certain systemic factors (method of conviction, county of conviction, identity of sentencing judge) and selected defendant characteristics (sex, age, race).
- The impact of habitual offender sentencing under the legislative guidelines on both proportionality and disparity.

The findings are summarized in Chapter Four and discussed in detail in Chapters Five through Fourteen.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Section III explores the policy implications of the research. Chapter Fifteen compares Michigan's guidelines to those of other jurisdictions. Chapter Sixteen discusses the change in required guidelines compliance from mandatory to advisory. Chapter Seventeen utilizes all that we have learned from the history, prior research, analysis of the guidelines structure, current empirical findings, and examples of other jurisdictions to propose changes that would improve the guidelines' capacity to achieve their objectives.

Whether mandatory or advisory, Michigan's sentencing guidelines still provide guideposts that are meant to increase proportionality and decrease disparity. With so much at stake, the value of an empirical analysis of Michigan's sentencing guidelines is evident. But such a study is not only a good thing to do, now is a good time to do it. Crime rates that are at a 60-year low,[4] a Michigan Department of Corrections budget that still exceeds $2 billion,[5] and a growing awareness that "get tough" policies can be counter-productive have moderated public attitudes. Legislators have exhibited a bipartisan willingness to address a number of controversial criminal justice issues.[6] It is hoped that the moment to consider so consequential an issue as the appropriate sentences for serious assaultive crimes has also arrived.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Section I. History and Context

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER ONE. THE EVOLUTION OF SENTENCING GUIDELINES

Until the mid-1970s, indeterminate sentencing was the norm all over America. It was based on the belief that sentences tailored to individuals could best achieve the "rehabilitative ideal." The thinking was that judges can evaluate each defendant's background and behavior to determine an appropriate minimum punishment and corrections officials can best determine when someone is sufficiently rehabilitated to return to society.

In its purest form, indeterminate sentencing is awash in discretion. Eventually, it fostered dissatisfaction from multiple perspectives. Some people were concerned about disparities among defendants and abuses of discretion in individual cases and sought a system with more consistency. Some thought defendants were not incarcerated for long enough. Some focused on the lack of accountability among decision-makers. Beginning in the mid-1970s, various reforms were adopted to varying degrees and in many combinations by jurisdictions across the country. Today, no two state sentencing and corrections schemes are exactly alike.

Michigan has retained the traditional indeterminate sentencing scheme it first adopted by state constitutional amendment in 1902.[1] Historically, the sheer length of the minimum sentence could not be reviewed by a higher court so long as it was not based on a constitutionally or statutorily impermissible factor and did not rise to the level of cruel or unusual punishment.[2] However, in 1972, the Michigan Supreme Court had before it "a plethora of cases involving sentences with a period of but 30 days between minimum and maximum."[3] In *People v. Tanner,* it concluded that such sentences violate the clear intent and purpose of the indeterminate sentence act by leaving virtually no time for the parole board to exercise its statutory discretion. As a result, the Court adopted a rule prohibiting imposition of a minimum sentence that exceeds two-thirds of the maximum.[4] By the end of the decade, the Court was ready to tackle the problems of sentencing on a much larger scale.

A. The Michigan Felony Sentencing Project Creates a Vision

In April 1978, the Michigan State Court Administrative Office established the Michigan Felony Sentencing Project (MFSP). A 19-member Steering and Policy Committee oversaw a staff of four researchers who were supported by dozens of data collectors, consultants, and advisors. The staff analyzed about a quarter of all felony sentences imposed in Michigan in 1977—approximately 6,000 cases. Random sampling was performed in a manner that ensured a sufficient number of cases for analysis by crime type and geographic region. Just 16 months later, in July 1979, the MFSP published its report, *Sentencing in Michigan*.[5] This extremely thorough analysis and the resulting recommendations laid the groundwork for how Michigan's sentencing scheme has evolved ever since.

The MFSP report identified three core problems: disparity, lack of accountability and the diffusion of authority over the actual sentence to be served among prosecutors, judges and parole boards. Disparity comes in three types:

- Justifiable disparity that can be explained by differences in the facts of the offense or the background of the defendant.[6]
- Invidious disparity that is unjustified but explainable as derived from case-processing variables, such as the identity of the sentencing judge, conviction method or county of conviction or such personal characteristics as race or economic status[7]
- Arbitrary disparity cannot be systematically explained by empirical research.

Project staff tested the hypothesis that judges systematically consider defendant and offense variables, accord each a relevant weight and then add the weighted variables together to determine whether a defendant should be incarcerated (the In/Out decision) and, if so, for how long (the Length decision). They found that a large degree of variation in sentences within crime categories is not explained by

RECEIVED by MSC 4/12/2022 1:07:49 PM

offense and defendant variables. They tested multiple other variables in an effort to determine what accounted for the rest of the variation in judges' decisions.[8]

The researchers also designed grids with points for defendant variables on the horizontal axis and points for offense variables on the vertical axis. Each cell where these points intersected contained the range of minimum sentences that judges actually imposed in these cases. These ranges were enormously broad, such as 12 to 300 months.

While believing that neither the unexplained nor the unwarranted sources of disparity resulted from judges intentionally discriminating, the MSFP researchers nonetheless concluded the following: "In a state adhering to the same laws and professing the value of equal justice, this is not a situation which should be allowed to remain unchanged."[9] Their solution was to adopt sentencing guidelines based only on those offense and defendant variables that judges actually deem relevant.

Based on the empirical research, the MSFP Steering and Policy Committee said the Legislature should establish a broad-based and racially balanced sentencing guidelines commission to develop and promulgate guidelines. Until the Legislature acted, the Committee urged the Michigan Supreme Court to appoint its own guidelines commission. The Committee assumed a commission would monitor the use and effectiveness of its guidelines on a regular basis.

B. The Judiciary Adopts Voluntary Guidelines

The Supreme Court responded by appointing five trial judges to the Sentencing Guidelines Advisory Committee or SGAC. The SGAC developed an initial set of guidelines that were pilot tested, revised and ultimately implemented in 1983.

The judicial guidelines primarily focused on disparity. The SGAC did not purport to decide what the "right" sentences were; rather, the goal was to rein in outlier judges whose sentences were above or below the average for similar offenses and defendants.[10]

To this end, the SGAC decided to focus exclusively on the facts of the crime and the defendant's prior record. It rejected variables that:

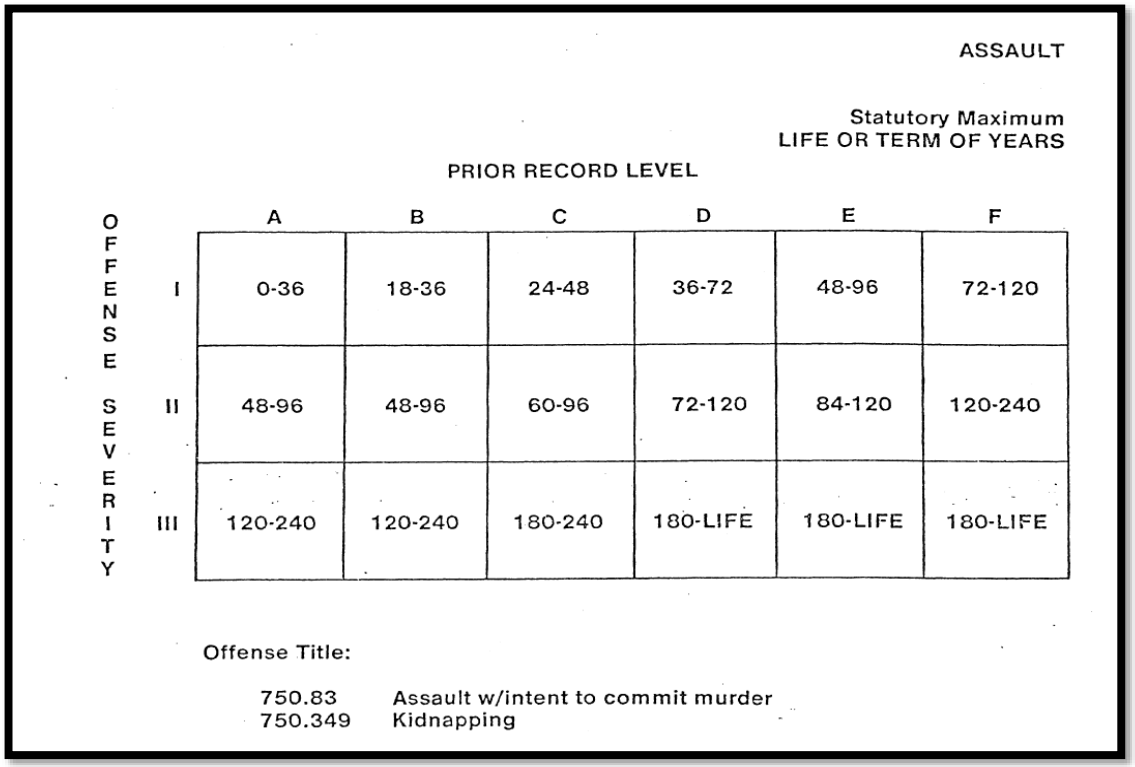
- Might increase racial disparity, such as demographic and socioeconomic considerations.
- Are not consistently either mitigating or aggravating, such as the prior relationship between the victim and defendant.
- Arise infrequently.
- Are not related to the goals of sentencing.
- Are not objective, such as remorse or the defendant's prospects for rehabilitation.[11]

The SGAC developed sets of prior record variables (PRVs) and offense variables (OVs). Each variable was scored according to its seriousness. The SGAC divided the most commonly occurring felonies into 11 crime groups.[12] Each crime group had multiple grids for more and less serious offenses as indicated by the statutory maximum penalty. Thus, for instance, within the assault group, five grids addressed offenses ranging from assault with intent to murder to attempted felonious assault.

Each grid had a horizontal axis divided into six levels for the PRV point scores, a vertical axis divided into three levels for the OV scores and cells containing recommended ranges of months within which the minimum sentence should be selected. The intersection of the OV and PRV levels identified the appropriate cell for the defendant's sentence. The lowest number in the cell—often referred to as the minimum-minimum or "min-min"—indicated the smallest number of months of incarceration that should be imposed. The highest number in the cell, often referred to as the maximum-minimum or "max-min," indicated the greatest number of months that should be imposed. For example, the grid for assault with intent to murder looked like this.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Figure 1. Assault Grid, First Judicial Guidelines



ASSAULT

Statutory Maximum
LIFE OR TERM OF YEARS

PRIOR RECORD LEVEL

OFFENSE SEVERITY	A	B	C	D	E	F
I	0-36	18-36	24-48	36-72	48-96	72-120
II	48-96	48-96	60-96	72-120	84-120	120-240
III	120-240	120-240	180-240	180-LIFE	180-LIFE	180-LIFE

Offense Title:

750.83 Assault w/intent to commit murder
750.349 Kidnapping

The recommended ranges were deliberately very broad, especially for the most serious offenses, leaving judges with a great deal of discretion. For instance, as Figure 1 shows, for someone being sentenced for assault with intent to murder who scored at the midpoint for both prior record level (Level D) and offense severity (Level II), the 72–120 month range was four years. Even so, the judicial guidelines were not mandatory. Judges were required to score them but they were not required to comply with their recommendations. When judges chose to depart from the guidelines, they simply had to complete a form explaining why their sentence was above or below the range.

The judicial guidelines were limited in their application, covering fewer than 100 frequently occurring offenses.[13] They did not apply to crimes with statutorily mandated sentences nor to defendants being sentenced for probation violations or as habitual offenders.

The SGAC made several important policy decisions. One was to adopt "real offense" sentencing. While the applicable sentencing grid is determined by the statutory maximum penalty for the conviction offense, the offense severity score is to be as high as all available evidence will support. If the proven or admitted facts are not consistent with the conviction offense, the actual facts are to be scored. So, for instance:

- If a defendant charged with armed robbery pleaded guilty to unarmed robbery to limit the maximum sentence to 15 years but use of a weapon was acknowledged, points for the weapon would be scored.
- Points may be scored for the aggregate value of property stolen or damaged even if some of that value was attributable to uncharged offenses or offenses that were dismissed pursuant to a plea agreement. Thus charges dismissed in plea bargaining do not really go away; they can come back to enhance the defendant's sentence precisely *because* they did not result in conviction.

RECEIVED by MSC 4/12/2022 1:07:49 PM

- If someone claims to have been a victim of a similar offense committed by the defendant that was never charged, the sentencing judge need only determine the accuracy of these claims by a preponderance of the evidence in order to consider them.[14]

The second policy decision was to decide which prior convictions, if any, are too old to be scored. The SGAC created a "10-year gap rule" that takes into account not only the time since the prior offense was committed but also the time during which the defendant was incarcerated or under supervision for that prior offense. The rule prohibits scoring a felony, misdemeanor or juvenile adjudication that occurred 10 or more years before the commission of the sentencing offense so long as the defendant had been discharged from that conviction for 10 years or more before committing any subsequent offense. Thus, the rule captures not only every offense that was committed during the preceding 10 years but also every offense the defendant ever committed before having a "clean" 10 years after discharge from incarceration or supervision.

The rule is applied in chain-like fashion so that each prior counted becomes the end-point of a new 10-year period. Take, for instance, a person charged in 2018 who was discharged in 2010 from a sentence imposed in 1995 for which he served 13 years in prison and two years on parole. If the person had a felony conviction in 1984, for which he received two years' probation (discharging in 1986) and a juvenile adjudication in 1975, the 1995, 1984 and 1975 offenses would all be scored as prior convictions even though they occurred 23, 34 and 43 years before the newest offense, respectively.[15]

Finally, the SGAC initially chose to reduce the guidelines score if one of three mitigating circumstances that were not legally sufficient to constitute a complete defense existed at the time of the crime:

1. Avoiding Harm (use of reasonable force to avoid harm to person or property)
2. Provocation/Passion
3. Mistake/Inadvertence (misunderstanding or ignorance of a significant fact)[16]

The method used was to calculate the guidelines scores, then reduce either the offense severity or prior record level by one step.

The SGAC monitored compliance with the guidelines, assessed departure explanations and considered feedback from trial judges. Racial disparity declined below statistically significant levels. The rate of compliance with the guidelines overall was roughly 80%. However, for the major violent crime groups (assault, robbery and criminal sexual conduct) the compliance rate was more like 60%. In many instances, people in the same guidelines cell were dissimilar in terms of the factors judges considered most important. [17]

The SGAC concluded that the guidelines needed to better capture the reasoning process of the sentencing judge. As a result, it made three fundamental changes.[18] It restructured the scoring system so that the points awarded better captured the order and magnitude of each category within a given variable and in relation to other variables. It reconfigured the grid structure from 3x6 to 4x4 on the rationale that four levels of prior record were sufficient to distinguish among defendants but that a fourth level of offense severity was needed to separate the various degrees of culpability. The SGAC believed that these changes yielded grid cells that contained people who were, "in fact, similar in terms of those factors that are most salient to the sentencing decision."[19] Third, the SGAC revised the recommended sentence range. It also omitted any scoring of mitigating circumstances and added Prior Record Variable 7, which awarded points for subsequent and concurrent felony convictions. The express goal was not to make sentences more or less harsh but "to continue the effort . . . to reduce sentencing disparity" and to ensure that the recommended sentences reflected the "going rate."[20] Nonetheless, the outcome was to increase the recommended ranges for some crimes, particularly CSC, substantially.

The second edition of the judicial guidelines became effective on October 1, 1988. It was never updated and so did not apply to any crimes created by the Legislature after 1988. Scoring the guidelines continued to be mandatory but actually using them to set sentences remained voluntary. Yet, by 1995, the

RECEIVED by MSC 4/12/2022 1:07:49 PM

overall compliance rate was 88%, and the rates for the major violent crimes had also increased,[21] not because more judges were following the guidelines but because the guidelines more closely followed the judges.

C. *Milbourn* Sets a "Proportionality" Standard for Appellate Review

Although compliance with the sentencing guidelines was not required, there was a nascent willingness to address sentences that were excessively long. The Michigan Supreme Court had to decide what standard to apply when reviewing sentences on appeal. In 1983, the Court held in *People v. Coles* that appellate courts could review sentences imposed by trial courts for an abuse of discretion that "shocked the conscience."[22] However, the *Coles* standard proved unworkable.

In December 1984, a young man named Kevin Milbourn broke into the apartment he had previously shared with his former girlfriend and trashed it, causing several hundred dollars' worth of damage. The break-in was part of a series of events over several weeks, including confrontations and damage to the victim's car, that arose from the difficulty Mr. Milbourn was having with the break-up. Milbourn had no prior record, and the sentencing guidelines recommended a minimum prison term between 12 and 30 months. The trial judge imposed a prison sentence of 10 to 15 years. The Court of Appeals found that its conscience was not shocked and affirmed the sentence.

When it considered Mr. Milbourn's case in 1990,[23] the Michigan Supreme Court noted that in the thousands of sentences reviewed since *Coles* was decided, the Court of Appeals had been shocked by only a handful. On a number of occasions, Court of Appeals judges had expressed the desire for further guidance on how the *Coles* standard should be applied. The primary difficulty, the Court noted, was with *Coles*'s complete subjectivity.

The *Milbourn* majority found that Michigan's legislative scheme reflects the "principle of proportionality"; that is, the concept that the most severe punishments should be reserved for the most serious crimes and that people who have prior records should be treated more harshly than those who do not. It noted that by providing a range of punishments for most felonies, the Legislature also recognized that any given instance of a particular crime may be more or less serious. The purpose of discretionary sentencing, the majority concluded, is not "to accommodate subjective, philosophical differences among judges." Rather, the task of the judiciary is to determine "where, on the continuum from the least to the most serious situations, an individual case falls" and to impose sentence accordingly.[24]

For assistance in making that determination, *Milbourn* directed judges to rely on the sentencing guidelines. Although compliance with the guidelines remained voluntary, the gradation of recommended ranges reflected the consensus of most judges on the appropriate penalty for each combination of sentencing factors. Departures from the guidelines should be based on reasons that were not already covered by guidelines variables or were not scored points sufficient for the circumstances. Departures not supported by such reasons "would alert the appellate court to the possibility that the sentence under review was disproportionate.[25]

The majority noted that following the guidelines would not unduly restrict judicial discretion, as the recommended sentence ranges were "generally quite broad."[26] Even where some departure from the recommended range was warranted, the extent of the departure was also subject to review for proportionality. The majority concluded that while Mr. Milbourn's behavior might warrant some departure from the guidelines, imposing the maximum penalty allowed for breaking and entering was disproportionate and therefore an abuse of discretion.

Milbourn brought intensive scrutiny to trial court sentencing practices. As will be discussed in more detail in Chapter Sixteen, the results of that scrutiny varied widely. While resentencing was ordered in many cases, the Supreme Court itself upheld extreme upward departures from the guidelines range in others. The Court of Appeals was left without clear guidance and various subjective grounds for departing evolved.

The extent to which departures are allowed presents an enormous challenge to any guidelines system. No matter how carefully designed a scoring system is, inevitably there will be unique cases that

RECEIVED by MSC 4/12/2022 1:07:49 PM

guidelines cannot easily accommodate. Judges understandably want the ability to tailor sentences to the facts of such cases. Yet if individual sentencing judges are routinely allowed to depart from guidelines recommendations that they believe do not fit the crime or the defendant, the goals of proportionality and consistency will be undermined. Trying to simultaneously preserve and restrain judicial discretion is a high-wire act that requires a lot of skill.

D. The Legislature Enacts Mandatory Guidelines

In 1998 the Michigan Legislature enacted its own version of sentencing guidelines. Four years earlier, MCL 769.32 established a sentencing commission that was quite different from the small group of judges that constituted the SGAC. The legislative commission had 19 members, including eight legislators and representatives of various stakeholders, such as the defense, prosecution, corrections and law enforcement but only had slots for two judges.[27] It was structured to ensure that compromises would be reached before legislators had to vote up or down on the entire package.

The Sentencing Commission did its work during the height of the "tough on crime" movement. Although Michigan's overall "index" crime rate has declined almost steadily from 1981 to the present, this decrease largely reflects the steady drop in property crimes that begin in that year.[28] The decline in violent crime lagged behind. The violent crime rate peaked in 1986, declined for three years, went up again in 1990 and peaked again in 1991, before declining steadily after 1993.[29]

Thus the perception in the 1990s was that harsh responses by the criminal justice system were necessary to stem the tide of assaultive offenses. These responses included a decline in parole approval rates, an increase in parole revocations for "technical" violations of parole conditions, the automatic waiver to adult court of juveniles 14 and older who committed serious assaultive offenses, and the construction of the Michigan Youth Correctional Facility, a private prison to house the wave of juvenile "super-predators" that ultimately never materialized. Most notably, passage of the guidelines legislation was conditioned on simultaneous passage of truth-in-sentencing legislation that was designed to lengthen the time prisoners actually serve by eliminating sentence reductions for good behavior and prohibiting placement in halfway houses before the entire minimum sentence was served.

The Legislature charged the commission with developing sentencing guidelines that accomplish multiple goals; these included public protection, reserving the most severe penalties for crimes of violence against people, identifying cases in which non-prison sanctions are appropriate and accounting for state and local correctional resources. MCL 769.33 (e) also said the sentencing guidelines should:

> (iii) Be proportionate to the seriousness of the offense and the offender's prior criminal record.
> (iv) Reduce sentencing disparities based on factors other than offense characteristics and offender characteristics and ensure that offenders with similar offense and offender characteristics receive substantially similar sentences.

1. Structural Differences

Structurally, the legislative sentencing guidelines are broadly similar to the judicial guidelines. Offense and prior record variables are scored, and the intersection of these scores on a grid determines the recommended sentencing range. Many of the variables are similar in both sets of guidelines. Offense variable scoring is still based on the "real" offense. There is still no scoring of mitigating circumstances. The "10-year gap" rule for counting prior convictions was retained. However, there are many significant differences:

- The legislative guidelines are policy-based. Recommended sentences are designed to implement deliberate choices about who should and should not go to prison and how long prison sentences should be, not simply reflect existing judicial practices.

RECEIVED by MSC 4/12/2022 1:07:49 PM

  - Penalties for less serious offenses committed by people with less serious criminal records were designated as "intermediate" or community-based sanctions, which may run the gamut from jail or residential treatment to probation, outpatient programs, community service, tether and fines. [30]
  - Prison sentences for more serious crimes and defendants with longer prior records were lengthened.
- The legislative guidelines were mandatory. Until the Michigan Supreme Court decided *People v. Lockridge* in 2015,[31] judges had to sentence within the recommended ranges unless they placed on the record substantial and compelling reasons for departing from the guidelines. MCL 769.34 (3). The reasons were subject to appellate review.
  - Under *Milbourn,* reviewing judges focused on the ultimate propriety of the sentence compared to the offense and the defendant, using the judicial guidelines to inform that assessment. The statutory scheme focused on enforcing the guidelines themselves as a means of promoting proportional sentences.
  - Substantial and compelling reasons had to:
    - be "objective and verifiable,"
    - "'keenly' or 'irresistibly' grab our attention,"
    - be "of 'considerable worth' in deciding the length of a sentence" and
    - exist "only in exceptional cases."[32]
- The legislative guidelines apply to all offenses except the handful with mandatory penalties that leave judges with no discretion to exercise.
- The legislative guidelines apply to habitual offenders.
- The legislative guidelines did not initially apply to probation violators but do now as the result of a 2005 Michigan Supreme Court decision.[33]
- The legislative guidelines contain "straddle cells," which leave complete discretion to trial judges in more than ten thousand cases annually as to whether to impose a prison sentence or an "intermediate" (community-based) sanction.[34]
- The legislative guidelines are structured somewhat differently.
  - The judicial guidelines grouped offenses into crime groups based on the nature of the offense, (e.g., assault, burglary, CSC, drugs, homicide). Within each crime group were separate grids depending on the seriousness of the offense, with seriousness defined by the maximum penalty, (e.g., armed robbery (life), unarmed robbery (15 years), larceny from a person (10 years)). The grids all had four rows for offense level and four columns for prior record level, for a total of 16 minimum sentence ranges. On every grid, some ranges were redundant. For instance, on the criminal sexual conduct grid, the 120–300 month range could be reached through three different combinations of offense and prior record levels.
  - The legislative guidelines divide all felonies into nine crime classifications based roughly on the statutory maximum term. Second-degree murder is its own class (M2). Otherwise, very different offenses with similar maximums are in the same class. Class A is for life-maximum offenses other than murder, including assault with intent to murder, first-degree criminal sexual conduct and armed robbery, even though they involve very different behaviors and cause very different kinds of harm.[35]

    Each crime class has its own grid that displays the recommended ranges for the minimum sentence, depending on the offense and prior record levels. The recommended ranges increase as the crime class increases. For instance, if the defendant had the lowest possible score for both offense and prior record, the range for the Class A offense of assault with intent to murder would be 21 to 35 months in

RECEIVED by MSC 4/12/2022 1:07:49 PM

prison while the range for the Class F offense of felonious assault would be 0 to 3 months with a presumptive intermediate (non-prison) sanction.

All of the grids have six prior record levels. Grids A–E also have six offense levels, for a total of 36 minimum sentence ranges per grid. The F grid has four offense levels; the M2, G and H grids have three. The increased number of ranges resulted in more redundancy. [The guidelines grids for M2 and Class A appear in Chapter Eight.]

The legislative guidelines reduce the crime classification if the offense was an attempt. If the attempted offense falls in classes A, B, C or D, the offense is scored on the E grid. If the attempted offense falls in classes E, F, or G, it is scored on the H grid.

The Sentencing Commission took two major steps that increased the lengths of prison sentences for the most serious offenses and defendants. One, it raised the recommended minimum sentence ranges. Two, for those defendants that prosecutors decided to treat as habitual offenders, it further broadened the ranges by raising the high end.

2. Raising the Minimums

It is difficult to compare the recommended sentences for each set of guidelines because the grid structure changed with each version. However, by looking at the corners of each grid, it is possible to at least compare the extremes (i.e. what the recommended ranges would be for the least and most serious cases). Table 1 shows the non-habitual offender sentencing ranges (in months) at the four corners of the grid for each set of guidelines for the four life-maximum offenses.

- The upper left corner is for the lowest score on both offense and prior record variables.
- The lower left corner is for the highest offense but lowest prior record scores.
- The upper right corner is for the highest prior record but lowest offense scores.
- The lower right corner is for the highest score on both offense and prior record variables.

Table 1. Comparison of Sentence Lengths under Judicial and Legislative Guidelines

Crime	Guidelines Period	Low Offense Low Priors	Hi Offense Low Priors	Low Offense Hi Priors	Hi Offense Hi Priors
Homicide	1st Judicial	12–84	120–life	72–216	240–life
	2nd Judicial	12–84	120–300/life	120–180	240–480/life
	Legislative	90–150	162–270/life	270–450/life	365–600/life
Assault with intent to murder	1st Judicial	0–36	120–240	72–120	180–life
	2nd Judicial	0–36	84–180	60–120	180–300/life
	Legislative	21–35	108–180	108–180	270–450/life
Criminal Sexual Conduct	1st Judicial	12–36	72–120	72–120	180–life
	2nd Judicial	12–72	96–240	72–180	240–480/life
	Legislative	21–35	108–180	108–180	270–450/life
Armed Robbery	1st Judicial	12–18	36–72	60–96	120–240
	2nd Judicial	0–36	36–96	36–120	120–300/life
	Legislative	21–35	108–180	108–180	270–450/life

RECEIVED by MSC 4/12/2022 1:07:49 PM

This very basic comparison reveals several points:

- The judicial guidelines treated each listed offense quite differently. In the second version, murder had the highest ranges, CSC was next, the assault ranges were substantially lower and the armed robbery ranges were the lowest by far. The legislative guidelines left murder on its own grid and raised the low ends of the murder ranges substantially. But they placed all the other life-maximum offenses on the A grid and thus treated them the same. The result was to push assault sentences up substantially and armed robbery sentences up even more while leaving CSC offenses, which already had very high ranges, less changed.
- Except for one corner for one offense,[36] the starting point of the legislative ranges (the minimum-minimum) was substantially higher than those of either set of judicial ranges for all of the most serious offenses. The extent of the increase varied depending on each offense's range under the second judicial scheme. For instance, the beginning of the range for someone with a low offense score for murder but multiple prior convictions increased by 12.5 years. For assault with intent to murder, the increase was four years; for CSC, it was three years and for armed robbery it was six.
- The cell ranges in every version of the guidelines were extremely broad. The legislative guidelines narrowed some and increased others. But except for the upper left corner of the "A" grid, the range in every cell was at least three years. Overall, sentences for people with similar offenses and similar prior records could vary by six, ten, 15 or even 20 years without judges having to depart from the recommended range; that is, enormous room for disparity, particularly for life-maximum offenses, was built into each incarnation of the guidelines that had reduction of disparity as their primary mission.

Both the breadth of the cells and the way in which the cell parameters changed from one set of guidelines to the next illustrate how the definition of proportionality is a moving target. While setting a baseline punishment for varying offenses is an inherently subjective process, it is still important to ask how the same set of facts can be "worth" six years one day and 15 years the next or, indeed, six years in one courtroom and 15 years in the courtroom next door.

3. Habitual Offenders

The habitual offender statutes, which have been around in some form since at least 1927,[37] were designed to give prosecutors and judges an extra tool in sentencing defendants with prior convictions. They were premised on the notion that each successive conviction reflects the defendant's unwillingness to change their criminal conduct. If a prosecutor decides to treat a defendant as a habitual offender, Michigan statutes provide a formula for increasing the maximum sentence beyond what the penal statutes normally allow.

- If the defendant receives a second offender enhancement, and the current offense carries a statutory maximum sentence less than life, the maximum may be raised by 50%.[38]
- If the defendant receives a third offender enhancement, and the current offense carries a maximum sentence less than life, the maximum sentence may be doubled.[39]
- If the defendant receives a fourth offender enhancement, and the current offense carries a maximum sentence less than five years, the maximum may be raised to 15 years. If the current offense carries a maximum of five or more years, the defendant may be sentenced to life or any term.[40]

RECEIVED by MSC 4/12/2022 1:07:49 PM

Of course, if the current offense already carries a maximum of life or any term, a habitual offender enhancement has little legal effect. In these cases, the court is already free to impose any maximum sentence it chooses and can take the defendant's entire prior record into account in making that decision. Whether habitualization has any practical effect in these cases depends on whether the court choses to give more weight to the defendant's criminal history because the prosecutor has chosen to utilize the enhancement.

Independent of sentencing guidelines, increasing the maximum sentence under the habitual offender statutes also has the collateral effect of allowing for longer minimum sentences. Since the *Tanner* rule described earlier prohibits the minimum sentence from exceeding two-thirds of the maximum sentence, lengthening the maximum creates room for the minimum to grow as well. Recall, however, that *Tanner* does not apply to life-maximum offenses so in those cases the court could impose any minimum it desired.

The judicial guidelines did not apply to defendants whom prosecutors treated as habitual offenders. Judges simply selected minimum sentences for these defendants as they always had without reference to the guidelines.

The legislative guidelines do not only cover people being sentenced as habitual offenders, but also increase the possible *minimum* sentence for those defendants by broadening the cell ranges. The high end of the range is raised by 25, 50, or 100%, depending on whether the person is being sentenced as a second, third or fourth offender. Thus, the judge can choose to impose a much longer minimum sentence while staying within the guidelines (i.e., without having to justify an upward departure).

Leveraging the habitual offender statute into a mechanism for raising both the minimum and maximum sentences raises multiple concerns that will be discussed in depth in Chapter Six. The bottom line is that the exclusive purpose of the habitual statutes had been to raise the statutory maximum. The sentencing guidelines are designed to structure the consideration of the defendant's prior record in selecting the minimum sentence. The more prior convictions the defendant has, the higher the prior record score and the longer the recommended minimum. Using the defendant's habitual offender status, which is also based on the defendant's prior convictions, to further enhance the recommended minimum sentence means that, for those defendants that prosecutors choose to treat as habitual offenders, prior criminal history is counted twice.

E. The Guidelines Are Not Assessed; Sentences Get Longer

The legislative guidelines were built on a raft of assumptions about how judges would actually use them. The enabling legislation assumed that the Sentencing Commission would review how the guidelines were working in practice and make periodic recommendations for modification. However, once the guidelines were submitted to the Legislature in 1997, the Commission never met again.[41] In 2002, the statutes that established the Commission and defined its tasks were repealed.[42] The assumptions that were used to develop the guidelines went untested. In the meantime, the Legislature proceeded to enact a host of provisions that undermined those assumptions, causing sentences to creep higher and higher. In her thorough and insightful analysis of trends from the adoption of the legislative guidelines in 1999 through 2013,[43] Anne Yantus explains the impact of these provisions.

1. Guidelines Amendments

Without a commission to provide context and consistency, the guidelines were periodically amended by the Legislature. Adjustments had to be made when new crimes were enacted or maximum penalties were changed. Increased maximums meant the offenses were moved to different sentencing grids with higher recommended minimums. In some cases, crimes were moved to higher grids without any change in the statutory penalty.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Yantus gives dozens of examples.[44] Two are particularly striking. When the guidelines took effect in 1999, first-degree child abuse was a Class C offense with a 15-year maximum. In 1999, the Legislature moved it to Class B. In 2012, the Legislature made child abuse a life-maximum offense and moved it to the A grid.

The crime of first-degree retail fraud was created with a two-year maximum in 1988 to distinguish shoplifters from people more appropriately charged with larceny in a building, which carries a four-year maximum. However, in 1998, the maximum for retail fraud was increased to five years. In 1999, retail fraud was placed on the H grid of the new sentencing guidelines. But in 2000, retail fraud was moved up to the E grid.

- As a result, 25% of people sentenced for first-degree retail fraud were pushed into presumptive prison sentences. Previously, 71% had been subject only to intermediate sanctions and 29% had fallen into straddle cells.[45]
- Moreover, retail fraud is now treated more seriously than larceny in a building, which still has a four-year maximum and remains on the G grid. It is also treated more seriously than felonious assault, which has a four-year maximum and is on the F grid.

Legislators also made changes that affected individual defendants. They tinkered with the scoring of various prior record and offense variables, repeatedly increasing point assessments. That resulted in moving affected defendants into higher guidelines ranges on the grids applicable to their crimes.

The effect of ad hoc legislative changes was three-fold. It undermined the comprehensive balancing of factors worked out by the Sentencing Commission and adopted by the Legislature itself when it enacted the guidelines. It altered the evaluation of proportionality for specific crimes by redetermining punishments without regard to the overall scheme. And, because virtually all the changes resulted in longer sentences and more incarceration, it increased the pressure on the prison system. Yantus concluded:

> A comprehensive review of Michigan's statutory sentencing guidelines from enactment in 1998 to 2013 reveals guidelines that have failed to evolve in a meaningful way. Instead, these guidelines reflect the politics of criminal law: a one-way ratchet that leads to ever-increasing penalties. [46]

2. Consecutive Sentences

Changes to the guidelines were not the only tool that legislators employed to lengthen sentences. Another was more consecutive sentences. Michigan's longstanding rule is that multiple sentences are to be served concurrently unless there is express statutory authority that requires or allows them to be consecutive. This rule holds whether the crimes arose from the same transaction or from different transactions that occurred before the defendant was arrested. There are two primary kinds of exceptions. One is for crimes by prisoners and others already under criminal justice supervision when they committed a new offense.[47] The other is for crimes specifically designated by the Legislature as requiring or permitting consecutive sentencing, such as possessing a firearm during the commission of a felony[48] and selected drug offenses.[49]

Yantus reports that, as of 1990, Michigan had 13 consecutive sentencing statutes, nine of which were mandatory while four gave trial judges the discretion to make sentences consecutive if they choose. By 2013, there were 42. Most of the 29 new provisions were discretionary. While many of the crimes involved were newly created, the penalties for some common longstanding offenses, like first-degree home invasion, first-degree CSC and resisting an officer were changed to allow for discretionary consecutive sentences.[50]

The cost of consecutive sentencing authority can be very high to both individuals and taxpayers. The Michigan Supreme Court has held that when a defendant receives consecutive sentences, so long as each one standing alone is proportional under *Milbourn,* their aggregate length will not be

RECEIVED by MSC 4/12/2022 1:07:49 PM

disproportionate. That is, each sentence is assessed separately; their cumulative effect need not be considered.[51] It has also adopted an interpretation of the CSC 1st statute that allows for consecutive sentencing on each count of CSC that occurred during the same criminal transaction (as opposed to just each different crime that occurred at the same time, such as CSC and robbery). Thus, where defendant Tommy Brown was convicted of three counts of first-degree CSC involving a minor that arose from the same incident, the Supreme Court upheld consecutive sentences of 40–60 years without regard to the fact doing so made the sentence 120 to 180 years.[52]

These extreme sentences are not limited to sexual assaults on children. Christopher Barron had four convictions that arose from a single 2000 offense: assault with intent to commit and conspiracy to commit armed robbery, carrying a concealed weapon and possessing a weapon while committing a felony. While serving his minimum prison term of 8.75 years, for those offenses, he was convicted of being a prisoner in possession of a weapon and received an additional sentence of 1.67 years. He was eventually paroled and was discharged from supervision in 2014.

Barron was then convicted of first-degree home invasion and armed robbery for forcing his way into the home of a Calhoun County couple in 2015. Armed with an authentic-looking BB gun, Barron had forced the man to lie down, struck him repeatedly and stolen about $1,500 in cash. The prosecutor treated Barron as a fourth habitual offender; the trial court imposed sentences of 25–60 years for the home invasion and 50–100 years for the armed robbery, to be served consecutively. With combined sentences of 75–160 years and no chance of reductions for good conduct, Barron will first become eligible for parole in 2091.[53]

These extraordinarily long consecutive sentences yield several observations.

1. Since they exceed the defendant's life span, they are obviously designed to avoid the possibility of release that a parolable life sentence would bring. Regardless of the person's age, health or changed behavior, the parole board will have no authority to grant release, and the public will continue to pay the cost of incarceration and medical care.
2. The Legislature's decisions about which crimes should allow for consecutive sentences do not reflect any principled rationale. The list of 29 consecutive sentencing provisions added from 1994–2013 ranges from impersonating a firefighter to identity theft to first-degree CSC and carjacking.[54] Hundreds of other offenses, including many that are hard to distinguish from those 29, still require concurrent sentences. Thus the changes appear to be ad hoc responses to situations that grabbed a legislator's attention.
3. Ad hoc decisions regarding consecutive sentences create anomalies. For instance, consecutive sentences are allowed for multiple counts of the same crime committed during a single incident, such as when a sexual assault victim was penetrated more than once during a single transaction. Yet the sentences for the same counts if committed at different times must be concurrent, such as when the victim was penetrated repeatedly in separate transactions over a period of time.[55]
4. Prosecutors control the number of offenses that are charged from a single criminal incident. In many cases where there are multiple victims or the chain of events violates multiple statutes, it is easy to increase the number of counts, thereby giving the prosecution more leverage in plea bargaining. Allowing multiple counts from the same transaction to be imposed consecutively without regard to their aggregate length allows prosecutors more room to increase the total time to be served.
5. Allowing proportionality to be assessed per count takes the focus off the real nature of the criminal behavior overall. If one person engaged in one criminal transaction, *Milbourn s*uggests that the total time to be served should be proportional to the entire transaction. Viewing the behavior as isolated component parts while ignoring their cumulative effect risks artificially inflating the aggregated sentence to the point of absurdity. Moreover, the guidelines that are used to inform review under *Milbourn* were constructed on the assumption

RECEIVED by MSC 4/12/2022 1:07:49 PM

that nearly all sentences would be concurrent and the longest one would control the minimum amount of time the defendant would actually serve.

In *People v. Norfleet*,[56] the Court of Appeals panel took an innovative approach to the issue of when and why discretionary consecutive sentences should be allowed. The defendant was sentenced as a fourth habitual offender on seven charges: three counts of delivery of less than 50 grams of heroin, one count of possession with intent to deliver less than 50 grams of heroin, one count of conspiracy to deliver less than 50 grams of heroin, one count of maintaining a drug house and one count of maintaining a drug vehicle. He was sentenced to consecutive terms of 134 months to 40 years for each of the first five counts plus concurrent terms of 46 months to 15 years for the last two, making him eligible for parole after 55 years.

The Court of Appeals recognized that the reasonableness of each of these sentences, which were at the high end of the recommended guidelines ranges, had to be reviewed on its own merits. However, the decision to impose the sentences consecutively was an exercise of discretion subject to review for abuse. The standard of review is whether the result is within the range of principled outcomes. In order to aid appellate review for abuse of discretion, the trial court must articulate its rationale for the imposition of each consecutive sentence; that is, it cannot treat the imposition of multiple consecutive sentences as a single act of discretion. The panel emphasized that Michigan historically considered consecutive sentences to be "strong medicine" and expressed the belief that requiring justification for each consecutive sentence will ensure that drastic deviations from the norm are reserved for appropriate situations.[57]

3. Mandatory Minimums

A second legislative technique that explicitly trumps the use of sentencing guidelines is the enactment of mandatory minimum sentences. Unless the penalty statute expressly allows for a departure, judges have no discretion to impose a lower minimum term. On the other hand, a mandatory minimum simply sets a floor. A judge can always choose to go higher.[58] Thus, the trial court's discretion to impose an individualized sentence is restricted for sentences below the mandatory penalty but not for sentences that exceed it.

Historically, Michigan's mandatory minimums were primarily aimed at drug offenses. However a massive overhaul of the drug laws in 2002 eliminated most of these. Yantus reports that seven common offenses currently have mandatory minimums. Among the harshest are two relatively recent enactments.[59]

In 2006, the Legislature added a 25-year mandatory minimum for first-degree CSC against a victim under the age of 13. Compared to the sentencing guidelines recommendation, this modification adds roughly 10 more years in prison for a person with no prior criminal record.[60]

In 2012, the Legislature added a 25-year mandatory minimum for selected people convicted of being fourth offenders.[61] The selection criteria are complex.

- The sentencing offense must be one defined by the statute as "serious."
  - The "serious" offenses are all assaultive although of varying degrees of severity.
  - They include, for instance, assault with intent to commit great bodily harm less than murder, which carries a 10-year maximum sentence, and assault with intent to commit unarmed robbery, which carries a 15-year maximum.
- At least one of the prior felonies must be among those defined as "listed."
  - The group of "listed" crimes is much longer.
  - Many are common and not all are assaultive.

So, for instance, a person charged with assault with intent to commit great bodily harm less than murder who has three very old prior convictions—two property offenses and the listed crime of CCW—must be sentenced to a minimum term of 25 years.

RECEIVED by MSC 4/12/2022 1:07:49 PM

As Yantus notes, many people will receive much longer sentences than they would have previously, even if the threat of these habitual offender minimums are used more as prosecutorial bargaining tools. The legislative analysis predicted the need for 1,600 more prison beds within 10 years and 7,300 beds after 25 years. The cost could be as high as $55.7 million per year after 10 years.[62]

4. A Lost Opportunity

In December 2014, the Legislature finally created a successor to the sentencing commission by establishing the Criminal Justice Policy Commission (CJPC).[63] Chief among the CJPC's broad array of tasks was evaluating the operation of the guidelines and recommending whatever modifications may be necessary to more effectively meet the Legislature's stated goals. The Legislature again specified a list of goals that modifications to the guidelines were to accomplish. The list was similar to that in the 1994 statute and exactly repeated the sections on proportionality and disparity quoted above.

However, the new statute went further. MCL 769. 33a permitted the commission to recommend changes not only to the sentencing guidelines but "to any law, administrative rule, or policy that affects sentencing or the use and length of incarceration." It also specified 10 policies that recommended modifications were to reflect. These policies emphasized such values as proportionality, equity, consistency, utilizing least restrictive means, the rational use of resources, reliance on research, transparency and the exercise of "judicial discretion to individualize sentences within a framework of law." Overall, the statute not only mandated criteria for the construction of the sentencing guidelines, it laid out the Legislature's broad vision for an integrated scheme of sentencing and corrections that would reduce reliance on incarceration to the extent that is conducive with public safety.

The CJPC statute contained a four-year sunset provision. Although the sunset was extended for nine months, the commission was ultimately allowed to die on September 30, 2019. The only aspect of the sentencing guidelines that it had examined in depth was the operation of the straddle cells.

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER TWO. CHANGING VIEWS OF HOW LONG IS LONG ENOUGH

While the current research specifically examines the impact of several versions of sentencing guidelines on sentences for four life-maximum offenses, it is useful to place the findings in a broader context. MDOC Annual Statistical Reports for the last 50 years provide a wealth of information about how much our concept of proportional punishment has changed. There was a time when 10 years in prison was considered severe punishment. In the 1970s, a large majority of everyone sentenced for second-degree murder, first-degree criminal sexual conduct, and armed robbery received a sentence that made them eligible for parole in 10 years or less. Less than 5% had minimums greater than 20 years.

People sentenced in the 2010s faced a very different reality. Through 2017, the percentage eligible for parole after 10 years had declined by 13 points for armed robbery, 57 points for first-degree criminal sexual conduct, and 62 points for second-degree murder. Conversely, the proportion of people required to serve more than 20 years for CSC had increased ten-fold, and the proportion required to serve more than 20 years for second-degree murder was nearly 15 times greater.[1]

Moreover, these changes must be viewed in the context of what judges well knew about the availability of sentence reduction credits. Under the progressive system of good time in effect until 1978, a 40-year sentence could be served in 16 years. Under the disciplinary credit system in effect from 1982 to 1998, a 40-year sentence could be served in 32.5 years. The minimums imposed after 1998 have no potential reductions for credit of any kind.

Notably, this trend toward ever-longer sentences runs counter to the norms of other countries that still view 10 years as an exceptionally long time to spend in prison. The 44 member states with data reported by the Council of Europe in 2018 collectively housed 1,229,385 prisoners; of these, only 1.5% were serving sentences of 20 years or more, and 1.2% were sentenced to life.[2] Many Latin American countries have banned any form of life imprisonment, as have nine member nations of the Council of Europe.[3]

A. 25-Year Trends and Mass Incarceration

1. The Michigan Evidence

From 1993 to 2017, index crimes plummeted by more than 56%, from 505,497 to 220,640.[4] The proportion of felony defendants sentenced to prison—as opposed to some other sanction—declined dramatically, from 29.3% in 1993, to 19.7% in 2017. Yet the Michigan prisoner population in 2017 (39,666) was 2.9% bigger than it was in 1993 (38,942). That failure of the prison population to decline can be explained by the ever-growing number of people serving very long sentences.

RECEIVED by MSC 4/12/2022 1:07:49 PM

From 1993 to 2017, the proportion of prison commitments with minimum sentences longer than 20 years more than doubled from 3.5% of the total to 7.4%.

The increase in very long minimum sentences also reflects a decline in the imposition of parolable life sentences, which dropped from 38 in 1993 to four in 2017. It appears that judges shifted from sentences that allow parole eligibility in 15 years to "big numbers" (i.e., minimums of 25 years or more).

The increases for specific life-maximum crimes were dramatic. The average sentence for second-degree murder rose by 29.0%, from 19 to 24.5 years. The average sentence for armed robbery grew by 32.3%, from 6.5 to 8.6 years. The average sentence for assault with intent to commit murder grew by "only" 9.3 %, from 12.0 to 13.1 years.[5]

More and more people serving very long sentences have changed the composition of the prison population.

- In 1993, more than 8,600 prisoners (including those serving life terms) had sentences that required them to serve more than 10 calendar years. That was 22.1% of the entire prisoner population. In 2017, more than 16,500 people—41.7% of the population—were serving sentences with effective minimums of more than 10 years.

- The number of people serving indeterminate terms (that is, non-lifers) with minimum sentences of more than 20 years grew from 1,622 (4.2%) to 4,843 (12.2%).

- In 2017, nearly 8,000 people—20% of the prisoner population—were serving either life terms or minimum sentences that exceeded 25 years. Half of those were serving non-parolable life terms that would end only with commutation or death.

Life Sentences in Michigan

Non-parolable

Life without the possibility of parole is mandatory for first-degree murder and a very small number of other offenses. Release is only by death or commutation.

Parolable

All of the most serious offenses short of first-degree murder carry a penalty of life or any term. This allows the trial judge to either impose a term of life with the possibility of parole or an indeterminate sentence, for which the judge selects both the minimum and maximum terms.

Parolable lifers never earned good time or disciplinary credits even when these were available. Until 1992, under the "lifer law," they became eligible for release after serving 10 calendar years. Lifers sentenced since then must serve at least 15 calendar years. In practice, parole at first eligibility is rare. Hundreds of parolable lifers serve 25, 35, or 45 years before release, or are left to die in prison.

Lifer Law

Until the passage of the 1978 "truth in sentencing" ballot proposal, the lifer law created the possibility of parole for anyone who had served 10 calendar years. That is, it applied to people with very long minimum sentences, not just to parolable lifers. Since then, a person with a long minimum must serve every day imposed by the judge. Thus, in theory at least, a lifer can be paroled in 15 years while someone with a 40-year minimum must serve 40 years.

The increase in people serving very long minimums, combined with a decrease in the number serving short sentences, has caused a shift in the overall population. In 1993, 57% had minimum sentences of five years or less; in 2017, it was just 40%. Some prison sentences may have become shorter; others may have become community-based sanctions.

While the number of prisoners serving shorter minimums has declined substantially, the increases for the longer sentences are dramatic. For example, the number of people with minimums of two years or less decreased by 23%, while the number with minimums over 25 years (not including lifers) grew 193%.

RECEIVED by MSC 4/12/2022 1:07:49 PM

This effectively means that 2,188 people serving two or fewer years were replaced by 1,780 people serving more than 25 years. Depending on their actual sentences, to serve their minimum terms, the missing short-termers would have required fewer than 4,376 prison beds over the course of their incarceration, while the additional very long-termers will require well over 44,500 beds.

This trend is not unique to Michigan. Researchers at the Urban Institute examined data from 44 states to analyze growth in prisoners' length of stay from 2000 to 2014—the result of policy changes in the 1980s and 1990s.[6] The extent of the increase varies across states, reflecting differences in state-level decision-making. In most states the trend was driven primarily or entirely by an increase in time served for violent crimes.

In Michigan, nearly one in four people had been in prison for more than 10 years. The average time served by the top (longest-serving) 10% of the population was 10 years in 1989, 16.6 years in 2000, and 26 years in 2013—a 160% increase in 25 years.

In sum, much longer sentences for those convicted of assaultive offenses —sentences that (if imposed after 1998) cannot be reduced by any award of disciplinary credits—have helped push up the average length of stay for Michigan prisoners. In turn, the increased length of stay keeps the size of Michigan's prisoner population from falling much below current levels despite efforts to divert more people convicted of non-violent offenses to community-based sanctions, rationalize drug laws, increase parole rates and return fewer people to prison for probation and parole violations.

2. Criminologists Weigh In

Criminologists and reform advocates increasingly agree that the problem of mass incarceration cannot be solved without addressing the long sentences imposed on people convicted of assaultive and sex offenses. Marie Gottschalk argues against leaving the "worst of the worst" out of the discussion about criminal justice reform strategies because they typically serve the longest sentences and make up a large portion of imprisoned people.[7] Prescott, Pyle, and Starr maintain: "Policies that seek to shrink the expansive prison population while ignoring prisoners who have committed violent offenses will fail to address the core of the problem." [8] Mauer, Nellis, and Myers not only agree that mass incarceration cannot be remedied without addressing long sentences, but also argue that sentences longer than 20 years are both cruel and impractical, as they do nothing to improve public safety and they add to the social and fiscal costs of incarceration.[9]

According to John Pfaff, the standard narrative that mass incarceration is driven by drug and other non-violent offenses is wrong; in fact, most people are in prison for serious, violent offenses. Based on his analysis of the research, he concludes that relying on long sentences in the name of public safety is ineffective and potentially counter-productive.[10]

Ryan King and his colleagues at the Urban Institute examined this comparison empirically. Using 2012 data, they illustrated the relative impact in 15 states of various techniques for reducing prison populations. For Michigan, they calculated that if then-current trends continued to December 2021, the population would be 39,304. A 15% reduction in the length of stay by those convicted of violent offenses would reduce the projected baseline by 5%, while a 50% reduction in the admission of those convicted of drug offenses would reduce the population by 3%.[11]

Austin et al observe that grouping people under the label "violent offender" has a host of consequences, including more pretrial detention, less access to alternatives to incarceration, longer prison sentences, and less eligibility for in-prison programs, lower security levels, and good time credits. The authors argue that the label—shorthand for multiple types of crime—often does not reflect the defendant's actual behavior and is not associated with high rates of recidivism, as policymakers fear. Generally ignored is the fact that violence is situational and that people who commit violent offenses have themselves often been victims of or witnesses to violence.

The authors conclude that the label "violent offender" undermines the principle of parsimony (imposing the least severe punishment necessary to achieve the goals of sentencing), distorts proportionality, and fails as a tool to predict future violent behavior. They recommend the following:

RECEIVED by MSC 4/12/2022 1:07:49 PM

1. Policymakers should make pretrial, prison classification, and parole decisions without the "violent offender" label.
2. Shorten prison sentences and lengths of stay for people convicted of violent offenses.
3. Make significant and lasting investments in social policy for communities challenged by violence and provide trauma-informed care and restorative justice options.[12]

B. Consequences of Longer Sentences

The increasing length of the sentences imposed under the legislative guidelines puts into stark perspective questions about how much punishment is proportional to a given crime and what limits, if any, there should be. As a group, the people convicted of these crimes are the least sympathetic prisoners in the system. Whatever the reasons for their behavior, in most cases they have caused serious harm to other people. However, it is because these defendants are the most vulnerable to sentences based on emotion and public pressure that the guard rails ostensibly erected by sentencing guidelines are particularly important. A broad consensus that some amount of prison time is warranted does not mean that *any* amount of time is acceptable or that all the collateral consequences of very long sentences are worth the cost.

1. The Fiscal Costs

The most obvious cost of longer sentences is the additional prison beds that are required. Consider, for example, a person convicted of armed robbery. The MDOC reports that the average sentence for armed robbery was 6.1 years in 1991[13] and 9.2 years in 2020,[14] a difference of 3.1 years. The 2021 average per-prisoner cost is projected to be $40,139 a year.[15] Thus, in today's dollars, the cost of incarcerating just one person convicted of armed robbery grew by $124,431. Lengthening the average sentence of 100 people imprisoned for armed robbery costs taxpayers over $12 million. And this calculation does not begin to count even larger increases in the average sentences for other life-maximum offenses.

The annual per-prisoner cost includes medical care. As prisoners age, they face the same illnesses and disabilities as the rest of us. They get cancer and heart disease and kidney failure and Alzheimer's. Their vision, hearing and mobility decline. Caring for the medically frail has become an inevitable consequence of very long sentences.

The House Fiscal Agency (HFA) reports that for the 15-year period from 2006 to 2020, per-prisoner costs for health care increased by an average of nearly 7.0% annually. Total health care spending for the period rose from $262 million to $296 million.[16] HFA says a major factor in the increase is the aging of the prisoner population.

- In 2006, 39.1% of prisoners were over age 40; 14.4% were over age 50.
- By 2020, those percentages had increased to 48.0% over age 40 and 25.6% over age 50.
- In 2020, 9.6% of Michigan prisoners—more than 3,200 individuals—were age 60 or older.[17]

The MDOC has a hospice unit, a dialysis unit and specialized geriatric units. Our prisons, already known for the role they play in housing the mentally ill, are increasingly becoming nursing homes as well.[18]

RECEIVED by MSC 4/12/2022 1:07:49 PM

2. The Impact on Public Safety

The belief that incarceration into middle or old age is necessary to ensure public safety is simply not supported. A host of research has found that the majority of former prisoners do not commit new crimes, people serving for homicide and sex offenses, in particular, have very low recidivism rates, and long sentences have a margin of diminishing returns.

Contrary to images of prison as a revolving door, research shows that a substantial majority of former prisoners do not return; for example:

- In a study of people released in 17 states between 2000 and 2012, Rhodes et al. found that 68% did not return to prison for any reason.[19]
- In 2009, the Citizens Alliance on Prisons and Public Spending (CAPPS) published a study of nearly 77,000 Michigan prisoners sentenced to indeterminate terms after 1981 and released for the first time from 1986 through 1999. It found that roughly 63% did not return to prison within four years; of those that did, nearly 20% were returned for technical parole violations; only about 18% were returned with new sentences for new crimes.[20]
- The MDOC's 2019 Statistical Report has a three-year follow-up of people released in 2016. It shows overall returns at 26.7%—13.6% for technical violations and only 13.2% with new sentences.[21]

Criminologists have long discussed the concept of "aging out" of crime. The data show indisputably that older people commit less crime than younger people,[22] at least in part because they have matured and are less impulsive, because they have settled into jobs and families, and because they are less willing to endure criminal justice system involvement. Whatever the reasons, extremely long sentences are unnecessary to reduce the likelihood of re-offending. Nature takes care of that.

A factor in recidivism with particular relevance to the issue of long sentences is the nature of the conviction offense. The 2009 CAPPS study found that people convicted of economically motivated crimes were much more likely to be returned to prison within four years, either for a technical violation or a new sentence, than those whose crimes tended to be impulsive.[23] Success rates, defined as no return to prison within four years for any reason, were 80.1% for homicide and 77.6% for sex offenses.

In looking at the types of new offenses for which people returned, the data showed:

- While 17.6% of all the people studied received new sentences for some kind of new crime, only 4.5% of the new offenses were crimes against a person (defined as homicide, sex, assault, and robbery).
- Only 7.2% received new sentences for crimes of the same type as their original offense (e.g., burglars who were released and then convicted of another burglary). For the people most likely to receive long sentences, the proportions were even lower:
    - Robbery – 4.9%
    - Sex – 3.1%
    - Assault – 2.9%
    - Homicide – 0.5%

These figures are corroborated by similar findings in numerous jurisdictions.[24]

The trend appeared again in research conducted by CAPPS in 2014. In 2009, then-governor Granholm expanded the parole board and directed it to review thousands of prisoners who had served their minimum sentences but been denied parole. Not surprisingly, these decisions had often been based more on the nature of people's offenses than on their institutional records or risk assessment scores. As a result of this "continuance review" process, the number of paroles for people convicted of homicide and sex offenses more than doubled.

RECEIVED by MSC 4/12/2022 1:07:49 PM

CAPPS examined the re-offense rates of people serving for homicide and sex offenses who were paroled from 2007 through the first quarter of 2010.[25] *More than 99% did not return to prison within three years with a new sentence for a similar offense.*

- Of 820 people who had been serving for murder or manslaughter, just two (0.2%) returned to prison for a new homicide.
- Of 4,109 people who had been serving for a sex offense, 32 (0.8%) returned to prison for a new sex offense.

A five-year follow-up of 188 people paroled in Maryland as the result of a state Supreme Court decision found similar results. The average age of these parolees when released was 64. They had served an average of 39 years, primarily for murder and rape. Their recidivism rate was less than 1%.[26]

Prescott, Pyle, and Starr reviewed prior research studies, then analyzed the three-year reimprisonment rates of 1,175,096 people released from prison in New York and California from 1991 to 2013, with a primary focus on 23,982 people convicted of murder and non-negligent manslaughter.[27] They found:

- The overall re-incarceration rate for individuals released after an initial incarceration for murder or non-negligent manslaughter is the lowest of any offense category—4.4% compared to 9.9% for the entire group, regardless of offense type.[28]
- The murder/manslaughter group's re-incarceration rate for a new murder or non-negligent manslaughter offense is 1.3%, which is high relative to the 0.2% rate for all releasees but still very low in absolute terms.
- As with every offense category, re-incarceration rates for the murder/manslaughter group declined steadily with increasing age at release.
- Years served is also associated with a decreased risk of recidivism but that correlation could not be disentangled from older age as a causal factor.

The authors conclude that older people convicted of violent offenses, particularly homicide, pose very little risk to the public upon release. They stress that every study has found that the vast majority of people convicted of homicide (usually more than 99%) do not commit another homicide upon release.

Notably, the 2009 CAPPS study found no relationship between the amount of time served and success upon release. This finding accords with a large and diverse body of empirical research demonstrating that increased length of stay does not reduce recidivism.[29] That is, keeping people in prison longer does nothing to enhance public safety.

Imposing very long sentences in the name of public protection is both ironic and counter-productive. Such sentences are unsupported by actual data on recidivism and are most often imposed on the people who are the least likely to repeat their offenses in any event. They substantially hamper the ability of people to be productive and self-sufficient when they are eventually released. And they require taxpayers to spend tens of millions of dollars annually for the illusion of increased safety.

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER THREE. PRIOR RESEARCH ON MICHIGAN SENTENCING GUIDELINES

A handful of existing studies of Michigan sentencing are important for testing and explicating the current findings. The foundation was laid when Zalman and his colleagues first explored why guidelines were needed.

A. Michigan Felony Sentencing Project (1979)

Before the advent of any guidelines, the MFSP examined the extent to which offense and defendant characteristics explained both the In/Out decision and the length of sentences to incarceration. They then attempted to determine what other factors affected these sentencing decisions. Throughout their analysis the MFSP researchers grouped offenses into 10 crime types, ranging from homicide, assault, sex and robbery to drugs and other non-violent offenses. They formed these categories on the rationale that "judges and other participants in the sentencing process tend to mentally group crimes into general categories and may apply different policies to these groups."[1] That is, the nature of the crime is a factor to be controlled for in testing the impact of other factors. The MFSP researchers concluded:

- Depending on the crime type, the combination of offense and defendant characteristics explained anywhere from 34 to 72% of the variance in length among prison sentences, with most crime types falling between 47 and 56%.
- Defendant characteristics are far more important than offense variables in determining who is incarcerated. Offense characteristics are most important for determining how long sentences are for violent crimes.[2]
- In looking at the specific crimes of first-degree criminal sexual conduct, armed robbery and felonious assault, it appears that "judges vary their sentences over the entire range of possible sentences regardless of the relevant offense and offender variables."[3]
    - The variation is greatest where the life or any term penalty allows for the widest expanse of discretion to be exercised. "Thus, a wide expanse of sentencing discretion itself generates unexplained variation. This implies that in order to reduce unexplained variation of sentences, it is necessary to have some method which reduces the amount or structures the content of discretion."[4]
    - Sentences appear more coherent where the statutory maximum is low and the range of discretion is smaller.
- The sentencing patterns of individual judges vary widely. "It seems clear that the source of the greatest amount of variation in sentencing lies in the treatment of those cases with the most severe maximum penalties."[5]
- There is substantial geographic disparity in sentencing practices.
    - The researchers divided the counties into three strata based on the number of criminal dispositions in a county's circuit court, regardless of where in the state the county was located. The strata were:
        - I (metro) – Oakland, Wayne and Recorder's Court
        - II (urban) – Bay, Berrien, Calhoun, Genesee, Ingham, Jackson, Kalamazoo, Kent, Macomb, Monroe, Muskegon, Ottawa, Saginaw, St. Clair, and Washtenaw
        - III (rural) – all the remaining counties (except Keweenaw which reported no felonies in 1977)
    - Average sentences were longest in Stratum I.

RECEIVED by MSC 4/12/2022 1:07:49 PM

    - Outcomes are much more consistent in larger jurisdictions than in rural counties. The extent of the variance in sex offenses explained by offense and defendant characteristics was 79% for Stratum I, 61% for Stratum II, and 47% for Stratum III.[6]
  - Because "different decision rules are being used in the three strata . . . the sentence an offender receives is dependent, in part, upon where he is sentenced."[7]
- Taking In/Out and Length decisions together, "there is evidence of very distinct differences in the treatment of whites and non-whites." [8]
  - Non-whites receive longer sentences to a degree that is statistically significant for drugs, burglary, larceny, and especially sex offenses.[9]
  - The racial disparity may or may not indicate invidious discrimination. At a minimum, it requires understanding the relative impact of the relevant variables on each group.
- Criminal justice processing variables, including type of attorney, custodial status, and conviction method, all have an impact on sentences, particularly on the In/Out decision.
  - Since having appointed counsel and being unable to make bail are both proxies for lower socioeconomic status, and requesting a jury trial is a constitutional right, the researchers considered these to be further forms of invidious disparity.
  - However, they found a statistically significant increase in sentence length associated with conviction method only for homicide cases. A jury trial added 53 months and a bench trial added 26 months.[10]

The MFSP findings set the standards for measuring the success of sentencing guidelines. Several studies since have looked at aspects of how Michigan's guidelines have affected sentencing disparity in practice.

B. Abel F. Ekpunobi, "Judicial Decision Making Under Michigan Sentencing Guidelines" (1999)[11]

In his doctoral dissertation, Ekpunobi used both logistic and linear regression models to assess the effectiveness of the second judicial guidelines "in reducing or eliminating sentencing disparities – situations in which legally similar defendants receive dissimilar sentences." The data were taken from nearly 21,000 felony cases with sentences imposed in four counties from 1992 to 1997. Ekpunobi examined both legally relevant factors (prior record, nature and severity of offense, guidelines range, criminal justice status) and extra-legal variables (demographic characteristics of defendants, sentencing judges and counties). Although identified by pseudonyms, the counties were Wayne, Oakland, Saginaw, and Kalamazoo.[12]

Ekpunobi looked at both the likelihood of incarceration in either jail or prison (the In/Out decision) and the length of prison sentences. His findings included:

- Among similarly situated defendants, the odds of *receiving* a prison sentence were higher for non-whites in all four counties; however, the *length* of the sentence was significantly associated with race only in Saginaw. The odds of going to prison or jail are higher for men than for women, but women who were sentenced to prison actually had longer average terms than men in Wayne County.
- The identity of the county significantly affects the outcome for some counties but not all. "In general, the mid-size counties have significantly higher state imprisonment and overall incarceration levels than the larger counties."
  - The odds of receiving a prison sentence were higher in Saginaw and Kalamazoo than in Wayne, but there was not a significant difference for Oakland.
  - Oakland had a higher overall incarceration rate than Wayne because it used jail more often.

RECEIVED by MSC 4/12/2022 1:07:49 PM

    - When it came to the length of prison sentences, county was a statistically significant indicator for two counties. Average minimum sentences were lower in Oakland and higher in Saginaw.
- Extra-legal factors, along with criminal justice status (e.g., probation, parole), were more important in the In/Out decision than they were in determining the length of prison sentences. Prior felony convictions, sentencing guidelines, and offense type/severity were the major predictors of the average minimum prison term.[13]

Ekpunobi suggested that more research should be conducted regarding the impact of variables he was unable to examine, particularly the method of conviction and whether defense counsel was retained or appointed. He concluded that it appeared desirable to strengthen the guidelines "by providing more structure and improving the enforcement mechanism."

C. National Center for State Courts (2008)

The National Center for State Courts (NCSC) examined the results of very different sentencing guidelines systems in Michigan, Minnesota, and Virginia.[14] Minnesota has a mandatory system that affords judges the least discretion. Michigan's sentencing ranges are five to six times greater than Minnesota's.[15] Minnesota has a single grid with 77 cells compared to Michigan's nine grids with 258 cells. Virginia has a voluntary system with a very active sentencing commission but no appellate review of compliance. Instead of grids, it uses 15 worksheets with a list of factors keyed to the type of offense. Both Minnesota and Virginia recommend an explicit single sentence that is the midpoint of a range.[16]

The goal of the research was to test the extent to which each state's guidelines achieved the goals of predictability, proportionality, and non-discrimination. The questions asked were whether:

- Similarly situated defendants received similar sentences.
- The guidelines provided meaningful and proportional distinctions between more and less serious wrongdoing.
- There was evidence of discrimination in sentencing based on race, age, sex, geographic region, or conviction method.

1. Background

For Michigan, the researchers looked at 32,754 individuals sentenced in 2004. NCSC stated that the legislative guidelines were designed to trim extreme sentences but to preserve considerable judicial discretion and tie recommendations to past practice.

- The compliance rate is very high for grids D–H, which have relatively narrower cell ranges and lower statutory maximum sentences. Departure rates are much higher for the M2 and A grids, which have life or any term as the statutory maximum, even though the cell ranges are much broader and therefore already allow for substantial exercise of discretion.[17]

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 2. NCSC Findings: 2004 Departure Rates for Life-Maximum Crimes			
Grid	Depart Below	Comply	Depart Above
M2	34.6%	57.7%	7.7%
A	39.9%	55.6%	4.4%

- The authors noted that inconsistencies can arise not only from prohibited discrimination but also from differing judicial philosophies. Some judges may give more weight to moral culpability, apparent dangerousness, or rehabilitative prospects than to imposing similar sentences on people with similar guidelines scores.[18]

2. The Impact of Guidelines Elements on Sentence Length

The NCSC researchers examined both the In/Out decision and the length of prison sentences. They assessed the impact of multiple guidelines elements: offense severity, prior record level, offense level, grid cell type (intermediate sanction, straddle cell, presumptive prison), habitual offender status, and departure. Regarding sentence length, they found:

- The average sentence for second-degree murder was 260 months. The average sentence on the A grid was 102 months.[19]
- For Michigan, the guidelines elements accounted for 67% of the variation in sentence length. For Minnesota, the guidelines elements explained 86.1%, a figure the authors characterized as "extraordinary."[20]
- All the major elements of the guidelines framework were highly significant in predicting the sentence length, with conviction offense severity (i.e., crime class/grid) having the greatest impact.[21]
- When both guidelines elements and extra-guidelines factors were included in the analysis, the extent to which habitual offender status lengthened the predicted sentence was:
  - Habitual 2nd 7.4%
  - Habitual 3rd 24.7%
  - Habitual 4th 38.7%[22]
- After controlling for all elements of the guidelines, upward departures were 121% longer than the expected sentence; downward departures were 48.5% shorter than the expected sentence. The authors call the magnitude of Michigan's departures "quite extensive." By comparison, in Minnesota, an upward departure increased the expected sentence by 47%; downward departure decreased the expected sentence by 29.1%.[23]
- Redundant cells had very different estimated sentences. For example, six cells on the A grid have a minimum-minimum of 108 months. The estimated lengths of sentences in these cells were: A-VI = 138, B-V = 107, C-IV = 126, D-III = 145, E- II = 141, and F-I = 127.[24]
- On the A grid, PRV levels A and B yielded virtually identical estimated sentence lengths at each offense level.[25]

3. The Impact of Extra-Guidelines Factors

The NCSC researchers also examined two kinds of extra-guidelines factors. Defendant characteristics included age, race, and sex. Systemic factors included conviction method and region of the state. The NCSC findings on defendant characteristics were:

- Blacks were more likely to be in a straddle cell.[26]
- Women were less likely to receive a prison sentence; young black males were somewhat more likely to receive one.[27]

RECEIVED by MSC 4/12/2022 1:07:49 PM

- Blacks received a 6% decrease in the length of sentence.[28]
- The profile of black defendants was different from whites. They were:
  - More likely to be on the M2 and A grid.
  - More likely to score at PRV levels D, E, and F.
  - More likely to be sentenced as habitual offenders, with the extent of the difference increasing as the habitual level increases.
  - Less likely than whites to be represented in every OV level (60% of blacks were in OV I, the least serious offense level)
  - More likely to be sentenced in southeast Michigan (59% of all black defendants compared to 14% of whites resided there, and 60% of defendants sentenced there were black).[29]

Regarding systemic variables, the NCSC found:

- Sentences in SE Michigan (defined as the counties of Wayne, Oakland, Macomb, and Washtenaw) were 22% lower than sentences in the rest of the state.[30] The exception was for people convicted at trial in SE Michigan where the trial tax of 30% made the SE sentence roughly equal to that of someone convicted by plea outstate.[31]
- 86% of the difference between SE Michigan and outstate sentences was due to differences in the weight given by judges to the relevant variables, as opposed to the mean values of those variables.[32]
- Habitual offender status 2nd and 3rd were used less often in SE Michigan but habitual 4th was used twice as often. [33]
- Conviction at trial (including both bench and jury) increased the sentence by 17.7%.[34]
- The "trial tax" was two and one-half times higher in the large urban courts.[35]

4. NCSC Conclusions and Recommendations

The NCSC authors concluded:

- Although the purpose of state sentencing guidelines is to achieve statewide consistency, Michigan effectively has two sentencing "regimes," the circuit courts in the SE counties and those in the 53 other circuits.[36]
- What makes this geographic dichotomy possible is the magnitude of the guidelines ranges that allow judges to sentence quite differently without having to depart. "If the norms of the urban courts lead judges to look to the bottom of the ranges, while outstate judges look toward the top, there can be dramatic differences in sentencing outcomes . . . While there is little evidence of discrimination as usually conceived, consistency is being achieved in a manner that suggests discrimination of a different kind – rather than using suspect factors to set sentences, judges in Michigan have been free to apply different weights to the sentencing relevant factors."[37]
- While there is no evidence of direct racial discrimination, blacks may benefit from their concentration in SE Michigan, where sentences are lower (unless they go to trial), which could mask any type of systematic statewide race effect.[38]
- As author Kevin Reitz observed: "[T]he most obvious mediating levers of sentencing authority – the trial courts' departure policy and the intensity of appellate review – are not the only factors that matter. Other critical factors include the breadth or narrowness of guideline ranges and the simplicity or complexity of factual considerations that must be fed into guideline considerations."[39]

RECEIVED by MSC 4/12/2022 1:07:49 PM

The authors recommended:

- Combine PRV levels A and B.
- Reduce number of OV levels.
- Reduce number of straddle cells.[40]

The NCSC report highlights a tension inherent in the design of guidelines ranges. On the one hand, wide ranges are a source of disparity. On the other hand, broader ranges generally mean fewer departures, and the goal of a sentencing commission should be to minimize departures.[41] The report does not address the anomaly that Michigan's M2 and A grids are extremely broad and nonetheless have many departures.

Finally, the NCSC authors concluded that sentencing commissions should periodically assess the extent to which their guidelines are actually achieving their goals. They trenchantly remarked: "Established policies are no more vibrant and self-sustaining over time than they are self-executing at inception."[42]

D. The Council of State Governments (2014)

In its analysis of Michigan's sentencing practices, the Council of State Governments (CSG) reached multiple broad conclusions. Among them were:

- Application of Michigan's guidelines yields disparity in sentencing.[43]
- Minimum prison sentence ranges are wide, and sentences are distributed across the ranges and beyond.[44]
- Michigan's guidelines result in minimum sentences that are "all over the map."[45]

CSG compared all the non-habitual sentences imposed in 2012 to the minimum sentences required by the low end of the relevant guidelines ranges or minimum-minimum. It found the distribution to be:

Below min-min	15%
At min-min	12%
10–190% of min-min	35%
200–290% of min-min	15%
300–390% of min-min	6%
400% or more of min-min	17%

CSG also noted that people who fell into the same straddle cell received very different punishments,[46] from 4–4.5% of guidelines prison sentences involved consecutives,[47] and as we will discuss in Chapter Fourteen, it found wide and troubling disparities in the use of habitual offender sentences.

E. Criminal Justice Policy Commission (2019)

The CJPC examined the sentences of 18,841 people whose guidelines scores placed them in straddle cells from 2012 to 2017. It found substantial disparity in whether straddle cell defendants were sentenced to prison or to community-based sanctions depending in part on the county of conviction. Other factors affecting disparity included race, age, sex, employment status, representation by retained vs. appointed counsel, crime type and method of conviction (trial vs. plea).[48]

Before it was disbanded, the CJPC began examining the use of habitual offender sentences. Its preliminary findings will be discussed in Chapter Fourteen.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Section II. The Current Research

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER FOUR. A SUMMARY OF THE FINDINGS

Basic Background Information

The research examined 38,651 sentences, including over 21,000 that were imposed under advisory guidelines developed by the judiciary and nearly 17,000 imposed between 1999 and 2012 under the mandatory guidelines enacted by the Michigan Legislature.

The sentences were for four types of offenses that carry maximum sentences of life or any term: second-degree murder, assault with intent to murder, first-degree criminal sexual conduct, and armed robbery. They were imposed in 13 counties and included nearly 5,600 with habitual offender enhancements. The data are consistently analyzed by offense type and habitual offender status to avoid confounding the effects of these key variables.

The findings address the sheer length of sentences and disparities that may be caused by particular factors. The unit of comparison is the median sentence length. Bivariate analysis was used to determine whether the relationship between sentence length and individual variables or between two individual variables, such as sex and prior record, was statistically significant (i.e., that it did not occur by chance). Throughout the report, the word "significant" is used to indicate a high degree of certainty that a factor had an impact on the length of the sentence. For more information about the data and the methods of analysis, see Technical Appendix A.

Multiple regression analysis was used to simultaneously examine the various factors known to impact sentence length in order to understand the unique contribution of each one. Regression interprets the impact of multiple independent variables at once so the effect of each can be seen while controlling for the others. Regressions can be used for two purposes: to identify how much of the variance in sentence length individual factors explain and to estimate how much a particular minimum sentence would change if a particular variable were adjusted. Key results of the regression analysis appear throughout the report. For a complete summary, see Technical Appendix B.

To understand the findings, the basic structure of the sentencing guidelines must be recalled. Offenses are divided into nine groups, depending on their statutory maximum sentence. Each group has its own grid containing recommended minimum sentences. The recommended sentences get longer from grid to grid as the seriousness of the offenses increase. Murder has its own grid, M2. The other life-maximum offenses are on the A grid.

Defendants receive points for various aspects of their prior record. The points are grouped into Prior Record Variable (PRV) Levels A–F, which are displayed horizontally across the top of each grid. Defendants also receive points for various details of the offense, which are grouped in Offense Variable (OV) levels and displayed vertically down the side of the grid.

The grid is divided into cells, where the PRV and OV levels intersect. Each cell contains a range of months within which the judge has full discretion to select the minimum sentence. The low end of the range is called the "minimum-minimum" or "min-min." The high end of the range is called the "maximum-minimum" or "max-min." A sentence below the min-min is called a "downward departure." A sentence above the max-min is called an "upward departure." If the defendant is being sentenced as a second, third, or fourth habitual offender, the max-min can be increased by 25, 50, or 100%, respectively.

What follows are 21 key findings from the research and, for most, brief summaries of the facts that support each finding. Substantially more detail can be found in Chapters Five through Fifteen.

RECEIVED by MSC 4/12/2022 1:07:49 PM

1. **The legislative guidelines increased sentence lengths for life-maximum offenses.**

    A. Altogether, in today's dollars, the cost of increasing sentence lengths for life-maximum offenses was $1.2 billion over 14 years (1999–2012), with no evidence that longer sentences improve public safety.
        1. The median non-habitual sentence increased by 1.5 years, or 23%.The cost of this increase was $51 million per year or $720 million for the 14-year period of study.
        2. The median habitual sentence increased by five years, or 50%. The cost of this increase was $35 million per year or $486 million over 14 years.

    B. The difference in median sentence length between non-habitual and habitual sentences grew substantially.
        1. Under the second judicial guidelines the median habitual sentence for all offenses was 3.5 years longer than the median non-habitual sentence.
        2. Under the legislative guidelines the difference between non-habitual and habitual median sentences overall grew to seven years.

    C. The increase in length of median sentences varied by offense and habitual offender status.

    D. The increase in length of median sentence varied depending on the county of conviction and each county's practices under the former judicial guidelines. For example:
        1. The medians for non-habitual offender sentences went from four to seven years in Oakland County, seven to eight years in Wayne County, and eight to 10 years in Ingham and Genesee Counties.
        2. The medians for habitual offender sentences went from seven to twelve years in Oakland County, 12 to 18 years in Wayne County, and 15 to 20 years in Genesee County.

2. **The percentage of sentences with habitual offender enhancements increased by more than a quarter, from 13.8 to 17.5% of the total.**

3. **The number of sentences per defendant increased under the legislative guidelines.**

    A. The frequency of multiple sentences is important because of the various ways they can be used to make sentences longer.

    B. Among individuals without habitual offender enhancements, 81.3% received only one sentence during the second judicial period. That proportion dropped to 74% under the legislative guidelines.

    C. Among individuals with habitual offender enhancements, the proportion who received just one sentence dropped from 78.6% under the judicial guidelines to 65.1% under the legislative guidelines.

    D. The frequency of single sentences under the legislative guidelines differed greatly by offense type and conviction method.

    E. The number of multiple sentences imposed varied substantially by county of conviction. For example:
        1. A CSC defendant in Oakland was nearly 10 times more likely to receive three or more sentences than a CSC defendant in Kent.
        2. A CSC defendant in Saginaw was nearly eight times more likely to receive three or more sentences than a CSC defendant in neighboring Genesee.

4. **The number of parolable life sentences has declined over time as the number of long indeterminate terms has grown. After peaking at 96 in 1993, the number of parolable life sentences imposed for offenses committed and sentenced in 2007–2012 never exceeded 17.**

RECEIVED by MSC 4/12/2022 1:07:49 PM

5. **The four critical guidelines components—offense type, defendant's prior record, severity of the way in which the offense was committed and the defendant's habitual offender status—together explained 70.9% of the variance in sentence length overall.**

    A. The extent of the overall variance explained differed significantly by offense type, ranging from just 28.9% for murder to 67.9% for robbery.
    B. The extent of the variance explained by the guidelines components differed significantly when examined by county of conviction.
        1. Guidelines components explained just 65% of the variance for Wayne and Genesee Counties.
        2. Kent, Macomb Jackson and Ingham Counties were all between 70 and 75%.
        3. The extent of variance explained was more than 82% for the counties of Oakland, Kalamazoo, Washtenaw, Muskegon, Calhoun, and Berrien.

6. **Offense type is the single most important component, explaining 25.6% of the variance in sentence length overall.**

    A. Murder sentences, which are calculated on a separate grid to reflect the unique seriousness of the crime, are always the longest by far.
    B. Armed robbery is consistently viewed as the least serious of the life-maximum offenses and sentences for that offense are much lower.
    C. The extent to which offense type accounts for differences in sentence length depends on the county; for example, it explains 34% of the variance in Wayne but only 16% in Oakland.

7. **Although the scoring of the seven prior record variables is designed to move people into higher prior record levels with relative ease, the majority of life-maximum defendants who were not habitualized had criminal histories that were minimal or relatively modest.**

    A. Depending on the offense, 15–20% of the sentences fell into Level A, although that requires that the defendant received no points for any prior record variable. At the other extreme, only 5 to 9% fell into Level F.
    B. While at least 50% of the sentences for every offense type fell at Level C or below, the proportion ranged from just 50.5% for assault sentences to 70.9% for CSC sentences.
    C. A great majority of life-maximum sentences received zero points for high severity felonies or high or low severity juvenile adjudications.
    D. PRV 7, which adds 10 or 20 points for concurrent or subsequent convictions that are, by definition, not "prior" to the instant offense, is the single most consequential. PRV 7 allows for punishment for the sentencing offense to be increased on the basis of convictions that would not otherwise count. Overall, 47% of sentences were pushed into at least PRV Level C based on this single variable.
    E. The extent of the variance in sentence length explained by prior record level ranged from 19.2% for murder to 35.8% for robbery.
    F. The extent of the variance in sentence length explained by prior record level also differed by county, ranging from 17.6% for Wayne to 35.7% for Muskegon.
    G. How much sentence lengths increased with each increase in PRV level depended on the offense. Prior record level was roughly three times as important to the length of armed robbery sentences as it was to murder sentences.
    H. The impact of changes in PRV level varied by county. For example, in Genesee, sentence length increased by 23% for PRV Level B, and 176% for PRV Level F. In Kalamazoo, the increase at Level B was 57%, and at Level F, 271%.

RECEIVED by MSC 4/12/2022 1:07:49 PM

8. With only 20 points separating each of the six offense levels on the A grid and 50 points separating the three offense levels on the M2 grid, the offense variable scoring results in people easily reaching the 100-point cut-off for the highest level on either grid.

A. For murder, assault and CSC, OV Level I is virtually irrelevant.
    1. Of the three offense levels on the M2 grid, 53% of the murder sentences scored 100 or more points, placing them at Level III.
    2. Of the six offense levels on the A grid, 44.6% of the assault sentences scored at least 100 points, placing them at Level VI; 85.8% were at Levels IV, V, or VI.
    3. In the CSC group, 29% of the sentences scored at least 100 points, placing them at Level VI; 68% were at Levels IV, V, or VI.

B. For armed robbery the situation is reversed: 48% of the sentences scored at OV Levels I or II while only 6% were at Level VI.

C. The non-habitual sentences illustrate exactly how this occurs. For example:
    1. OV 6 scores a large number of points for the very intent to kill, which is an element of the offense for both murder and assault with intent to murder.
        a. Among murder sentences, 21.9% received 50 points and 49.8% received 25.
        b. Among assault with intent to murder sentences, 18.4% scored 50 points and 47.3% scored 25.
    2. OV 1 scores aggravated use of a weapon when, although not a legal element, that behavior is an overwhelmingly common means of committing murder, assault, and armed robbery.
        a. Over 61% of armed robbery sentences had 15 points for pointing a gun at or toward a victim or threatening a victim with a knife, although actual or purported *possession* of a weapon is an element of the offense.
        b. Nearly 75% of murder sentences and 80% of assault sentences had 25 points for discharging a firearm at a victim or for cutting or stabbing a victim.
    3. OV 12 scores felonious conduct not resulting in conviction that occurred within 24 hours of the sentencing offense.
        a. Using the policy of "real offense" sentencing, the sentence for the conviction offense can be increased based on uncharged conduct or charges dismissed in plea negotiations.
        b. Roughly 25% of sentences received five, 10, or 25 points for OV 12 except in the CSC group, where the proportion was 18%.
    4. OV 13 further enhance the penalties for defendants with a pattern of sexual assaults on one or more children younger than 13, for defendants with substantial prior criminal conduct, and for gang-related behavior. All crimes within a five-year period, including the sentencing offense, must be counted without regard to whether the offense resulted in a conviction.
        a. In the CSC group, 27.1% of the sentences scored 50 points; 26.6% scored 25 points.
        b. Awards of 25 points occurred in 13.3% of murder sentences, 25.3% of assault sentences, and 27.8% of robbery sentences.

D. The extent to which offense level explained variance in sentence length differed greatly from county to county.
    1. Wayne County at 11.9% and Ingham County at 12.2% were at the very low end.
    2. Oakland at 36.6% and Calhoun at 38.9% were at the other extreme.

E. OV level explained only 5.4% of the variance in sentence length for murder but between 24 and 29% for the other three offense types.

RECEIVED by MSC 4/12/2022 1:07:49 PM

F. Each increase in OV level had far more impact on CSC sentences than on any other offense type.

9. Under the legislative guidelines, the overall median for habitual offender sentences was 88% longer than the median for non-habitual sentences.

A. The median length of habitual offender sentences reflects larger prior record scores as well as their enhancement to habitual status.
B. When all other factors are held constant, we see the impact on sentence length of enhancement alone.
    1. Enhancement as a second habitual offender increased the sentence by 10%.
    2. Enhancement as a third habitual offender increased the sentence by 19%.
    3. Enhancement as a fourth offender increased the sentence by 25%.
    4. The effect of enhancements varied substantially by county.

10. Departures, which are essentially a rejection of guidelines recommendations, play a critical role in sentence length.

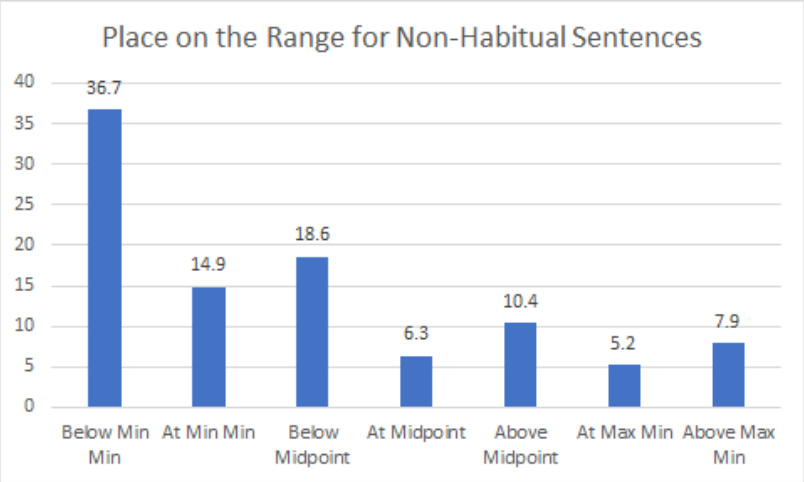
A. The overall compliance rate with the second judicial guidelines was 88%
B. The compliance rate for non-habitual sentences under the legislative guidelines on the M2 and A grids combined was 55.4%.
C. The rate of downward departures increased from the judicial guidelines for every offense, suggesting that judges often found the recommended ranges to be too severe. The proportion of downward departures was:
    1. Murder – 32.3%
    2. Assault – 38.6%
    3. CSC – 23.3%
    4. Robbery – 40.3%
D. The rate of upward departures also increased, albeit much more modestly. The proportion of upward departures was:
    1. Murder – 11.7%
    2. Assault – 10.7%
    3. CSC – 13.6%
    4. Robbery – 5.4%
E. Overall, downward departures decreased sentence length by 44%. Upward departures increased sentence length by 76%.
F. The extent of variance in sentence length explained by departures is roughly the opposite of the extent explained by the guidelines components. Although departures explain 18% of the variance overall, the rate differs greatly depending on offense type.
    1. For murder, departures explain fully 50% of the variance—almost twice as much as the guidelines components.
    2. For assault, CSC and armed robbery departures explain 30, 26, and 20% of the variance, respectively.
G. The extent of variance in sentence length explained by departures also differs substantially by county.
    1. For Wayne and Genesee, departures explain roughly 22% of the variance.
    2. For Kent, Macomb, Ingham, Jackson, and Saginaw, departures explain from 12.7 to 17.5% of the variance.
    3. For Oakland, Kalamazoo, Washtenaw, Berrien, Calhoun and Muskegon departures explain only 5 to 10% of the variance.

RECEIVED by MSC 4/12/2022 1:07:49 PM

11. The broad guidelines ranges promote disparity.

A. Some ranges are so broad that recommended minimum sentences for people with similar backgrounds and similar offenses can differ by seven, nine, 12, even 20 years. As a result, where within the applicable range sentences fall can be as important as upward and downward departures.

B. The graph shows that sentences were far from evenly distributed.



C. Adding the sentences at the min-min to the downward departures, and adding those at the max-min to the upward departures, demonstrates how many sentences for each offense type were relatively long and relatively short.
   1. For murder, 38.3% of the sentences were at the low end, and 19.7% were at the high end.
   2. For assault, 47.6% were at the low end, and 16.7% were at the high end.
   3. For CSC, 33.4 were at the low end, and 25.0% were at the high end.
   4. For robbery, 59.0% were at the low end and 8.5% were at the high end.

D. Judges' decisions about place on the range are not explained by PRV or OV levels.
   1. In every offense group, people with high prior record scores received downward departures and people with low prior record scores received upward departures.
   2. Having a low OV score does not increase the likelihood of a downward departure nor does having a high OV score increase the likelihood of an upward departure.

E. The disparity allowed by the guidelines can be best seen by comparing the sentences imposed on people with the same offense type whose PRV and OV scores combined to place them in the same guidelines range. The A grid has four cells that have a range of 135–225 months (11.25–18.75 years) with a midpoint of 15 years for non-habitual sentences. From 2003–2012, 223 people convicted of sex offenses fell into one of these cells. Of these:

- 43 received minimums below 11.25 years, the lowest being 3.5 years
- 12 were sentenced to exactly 11.25 years
- 46 received sentences above 11.25 and below 15 years,
- 39 were sentenced right at the midpoint to 15 years
- 28 received sentences greater than 15 but less than 18.75 years
- 21 were sentenced to exactly 18.75 years
- 34 received minimums above 18.75 years, the highest being 65 years

RECEIVED by MSC 4/12/2022 1:07:49 PM

12. The guidelines create the illusion of precision but so many cell ranges are repeated and overlapping that median sentences are inconsistent and unpredictable.

A. Over and over again, identical ranges produce widely different median sentences.
    1. The M2 grid has 18 cells but only eight different ranges. This mixture results in 12 different medians.
    2. The A grid has 36 cells but only 11 different ranges. These result in 26 different medians for assault, 25 medians for CSC, and 19 for robbery.
B. Not uncommonly, medians are at or below the low end of the range.

13. While habitual offender sentences are much longer than non-habitual sentences, fewer than half exceed the standard high end of the applicable range.

A. At 12.6%, the rate of downward departures was far lower for habitual than for non-habitual sentences. Nonetheless more than 10% of sentences at every habitual level were below the min-min.
B. At least 40% of the sentences at every habitual level were below the midpoint of their cell range—a little more than half the rate of non-habitual sentences. Much of the difference is attributable to the lower rate of downward departures.
C. Depending on the offense type the proportion of sentences above the standard max-min was 28.5% for second offender enhancements, 38.5% for third offenders, and 40.2% for fourth offenders.
D. Compared to non-habitual sentences, habitual offender enhancements resulted in one-third the rate of downward departures and four and a half times the rate of sentences above the high end of the guidelines. Between these extremes, the distribution along the range of habitual and non-habitual sentences did not differ substantially.
E. The extent to which sentences were above the standard max-min, regardless of habitual level, ranged from 26.3% for assault and robbery to 31.4% for CSC.
F. Habitual offender sentences that did exceed the standard max-min ran the gamut from only slightly above to well above even the 100% increase in the recommended range that the guidelines allow for fourth offenders.

14. Defendants with similar backgrounds and offenses received significantly different sentences depending on the county in which they are convicted.

A. County size appears to be more important than location. The median sentences imposed in the larger counties tended to be shorter than those imposed in smaller counties.
B. The data undermines some widely held assumptions about which counties are "tough" and which are "lenient." Wayne County, home to the City of Detroit, did not impose the lowest non-habitual sentences in any offense group. Oakland County sentences were in the low-to-middle range for every offense, despite the county's "law and order" reputation.
C. County differences in median sentence length for non-habitual sentences may reflect, in part, the extent to which defendants with high PRV levels have or have not been shifted into the habitual sentence group.
D. Murder sentences showed the greatest amount of consistency despite the fact that the M2 grid has the widest cell ranges by far. The median sentence in nine counties was 20 years; it ranged from 22.5 to 24 years in the other four counties.
E. CSC sentences showed the least consistency with a spread of nearly nine years between the lowest medians and the highest.

RECEIVED by MSC 4/12/2022 1:07:49 PM

F. Downward departure rates cannot be explained by county size or location. For example, rates for the four largest counties declined in increments of roughly 10 points:

- Wayne – 47.6%
- Kent – 36.0%
- Macomb – 26.4%
- Oakland – 16.5%

G. The obverse of departure rates is the extent to which counties comply with the sentencing guidelines.
    1. Wayne, the county with the largest number of cases by far, imposed sentences within the recommended range only 41.5% of the time. At 57.6, 59.5, and 60.8%, respectively, the counties with the next lowest compliance rates were Ingham, Kalamazoo, and Kent.
    2. At the other extreme, the smaller counties of Berrien, Muskegon, Calhoun, and Jackson imposed within-range sentences between 85 and 89% of the time.
    3. Notably, Oakland, at 81.1%, and Washtenaw, at 78.6%, behaved much more like the smaller counties than like their larger peers.

H. Combining the downward departures and minimum-minimums gives a picture of the extent to which counties imposed sentences at the low end.
    1. Five counties: 53.0–65.4%.
    2. Four counties: 38.5–45.3%.
    3. Four counties: 13.8–24.1%

I. The proportion of both low-end and high-end sentences for each offense varied across counties.
    1. The length of sentence overall was significantly longer in seven counties compared to Wayne, significantly shorter in three, and not much different in the remaining two. For instance, Genesee, Jackson, and Muskegon had sentences that were 7% longer, while Oakland's were 7% shorter.
    2. Similarly, Kent's sentences were 4% longer than Wayne's, while Macomb's were 4% shorter.

J. The disparities by county were far and away most pronounced for CSC sentences.
    1. Eight counties had sentences that were significantly different than Wayne's and all were longer.
    2. The greatest difference was with Genesee, which was 24% longer; the smallest was Kent, which was 5% longer.

K. Counties varied greatly in their rate of plea-based sentences for each type of offense.
    1. Plea rates for murder ranged from 54.3% in Macomb to 85.7% in Ingham.
    2. For assault, plea rates ranged from 50.0% in Jackson to 79.1% in Kent.
    3. For CSC, plea rates ranged from 62.9% in Ingham to 86.6% in Kalamazoo.
    4. For robbery, plea rates ranged from 79.5% in Jackson to 94.7% in Kent.

L. While one would expect a clear relationship between a large proportion of low-end sentences and a high rate of guilty pleas, in fact that relationship was inconsistent.
    1. Although nine counties all had median minimums of 20 years for murder, their rates of guilty pleas ranged from 54.3% for Macomb and 62.7% for Oakland to 85.7% for Ingham and 83.3% for Genesee.
    2. For the A grid offenses, the relationship between plea rates and low-end sentences was all over the map.

RECEIVED by MSC 4/12/2022 1:07:49 PM

15. **Although the available data were limited, it was apparent that judges in the same county imposed significantly different sentences for the same type of offense. Whether a defendant is made to serve 10 years in prison or 15 may well come down to whether their case was assigned to Judge A in Courtroom 1 or Judge B nextdoor in Courtroom 2.**

16. **The use of habitual offender enhancements varies significantly by county.**
    A. The percentage of eligible sentences that received habitual enhancement ranged from 13% in Wayne to 91.7% of eligible sentences in Muskegon and Saginaw Counties. The rate of application can be broken into three groups
        1. Kent, Macomb, Washtenaw, and Wayne Counties applied the enhancement to 13%–30% of eligible sentences.
        2. Berrien, Genesee, Ingham, Jackson, and Kalamazoo Counties applied the enhancement to 42%–52% of eligible sentences.
        3. Calhoun, Muskegon, Saginaw, and Oakland Counites applied the enhancement to more than two-thirds of eligible sentences.
    B. The rate of habitual enhancement differed by offense. Murder sentences were enhanced the least often, with 24.7% of eligible sentences receiving enhancement, and CSC sentences were the highest, with 44.7% receiving the enhancement.
    C. PRV level is significantly related to whether a sentence was enhanced or not. Generally, eligible sentences that had high PRV scores were more likely to receive habitual enhancement.
    D. The method of conviction is significantly related to the application of enhancement to eligible sentences.
        1. Approximately 80% of sentences that were not enhanced resulted from a plea, compared to 51.6% of sentences that were enhanced.
        2. Similarly, 16% of eligible but not enhanced sentences resulted from jury trial compared to 41.9% of sentences that were habitualized.
    E. There was no apparent pattern between how frequently a county applied the habitual enhancement and the size of the impact on sentence length.
        1. Kent and Wayne Counties applied the habitual enhancement to less than 25% of the eligible sentences, but when they did, the sentence length doubled between not enhanced and fourth habituals.
        2. Oakland County applied the habitual enhancement often but had the smallest increase in sentence length.

17. **The guidelines have reinforced disparities created by plea bargaining.**

    A. The rate of guilty pleas in life-maximum cases, which was 79.6% in 2019, is historically lower than it is for less serious offenses, which was 97.4%.
    B. Plea rates vary by offense type. For non-habitual sentences under the legislative guidelines the rates ranged from just 59% for assault to 85% for robbery.
    C. After the Legislature extended guidelines coverage to habitual offender sentences with the threat of increasing the high end of the recommended range by 25, 50, or 100%, guilty plea rates declined and jury trial rates increased dramatically.
        1. The decline in the plea rates and the increase in jury trials from the judicial period, when the guidelines did not apply to habitual sentences, to the legislative period was roughly 20 points for every offense type except murder.
        2. During the judicial period, depending on offense type, the plea rate for habitual sentences was just 0–7 points lower than for non-habitual sentences. Under the legislative guidelines, the plea rate for habitual sentences was 26–33 points lower than for non-habitual sentences.

RECEIVED by MSC 4/12/2022 1:07:49 PM

D. In both guidelines periods, defendants who were convicted by a jury received significantly longer sentences than those who pleaded guilty. The size of the trial tax, calculated as the percentage increase in sentence length for jury trials over pleas, varied by offense type.
    1. For non-habitual sentences, the trial tax during the judicial period ranged from 33.30% for murder to 150.0% for robbery.
    2. During the legislative period the extent of the trial tax on non-habitual sentences was roughly similar, ranging from 38.9% for murder to 118.0% for robbery.

E. Compared to the judicial period, under the legislative guidelines the trial tax on habitual offender sentences exploded. It grew from:
    1. 20% to 60% for murder
    2. 66.7% to 100% for assault
    3. 33.7% to 66.7% for CSC

    For robbery, the tax as a percentage increase declined from 100% to 77.8%, but the actual difference in sentence lengths remained substantial. During the judicial period, the median sentences were 7.5 years for pleas and 15 years for trials. Under the legislative guidelines, the median robbery sentence became 11.25 years for a plea and 20 years for a trial.

F. For non-habitual sentences, the rates of downward departures for guilty pleas vs. jury trials were dramatically different. Although relatively more modest, the differences in upward departures by conviction method were also substantial.

Percentage of Departures by Conviction Method for Non-Habitual Sentences				
Crime	Downward Departures		Upward Departures	
	Pleas	Jury Trials	Pleas	Jury Trials
Murder, second	39.8	9.2	10.0	17.9
Assault	55.0	9.9	6.5	17.1
CSC, first	28.3	9.2	9.0	23.9
Robbery	43.3	13.9	4.3	14.8

    1. Sentences right at the high end of the range after jury trials compared to pleas were three times as common for murder sentences, nine times as common for assault sentences, twice as common for CSC sentences and 3.5 times as common for robbery sentences.
    2. These differences were not explained by differences in the PRV or OV scores of people who pleaded guilty as opposed to those who went to trial.

G. Among habitual offender sentences, the differences in departures above the usual maximum-minimum between pleas and jury trials were even more extreme. For example:

RECEIVED by MSC 4/12/2022 1:07:49 PM

Percentage of Departures by Conviction Method for Habitual Sentences				
Crime	Plea		Jury	
	Downward	Upward	Downward	Upward
Murder – Hab 2	3.5	8.8	2.9	57.0
Assault – Hab 2	32.0	12.0	0.0	76.0
CSC – Hab 4	18.0	30.0	0.0	66.0
Robbery – Hab 4	19.0	18.0	1.0	60.0

H. The trial tax reflects not only the fact that the defendant went to trial but also all the factors associated with that decision, including the defendant's prior record, the details of the offense, habitual offender status, and the likelihood of a departure. When the regression analysis honed in solely on the impact of the conviction method while holding all other variables constant, it appeared that just the fact of having a jury trial added 18% to the sentence compared to a guilty plea.
    1. The amount of the increase varied by offense, ranging from 12% in murder cases to 21% when the conviction was for assault.
    2. The sentence increase for selecting a jury trial varied substantially by county.
        a. The fact of a jury trial added 11% in Ingham and Kent, 12% in Wayne, 15% in Kalamazoo and 19% in Washtenaw.
        b. A jury trial added 24–26% to the sentence in Genesee, Macomb, Muskegon, and Oakland, and 30% in Calhoun.
        c. The most striking county disparities were the lack of any significant increase for a jury trial in Berrien and Jackson and, at the other extreme, a 38% increase in sentence length for having a jury trial in Saginaw.

18. Although women received significantly shorter sentences than men, the reasons were not altogether clear.

A. Women accounted for nearly 11% of the murder sentences but less than 5% of other offense types.
B. For every offense, women's median sentences were shorter than those of men by varying percentages: murder - 17%, assault - 41%, CSC - 30%, robbery - 29%.
C. The difference in median sentence is attributable in part to women having significantly lower PRV levels than men.
D. Women's sentences were also significantly more likely to have resulted from guilty pleas.
E. Given that women's median sentences were from roughly two to five years shorter than men's, the differences may also result from a combination of factors not captured by the data.
F. Nonetheless, sex per se only makes a significant difference in assault sentences, which it lengthens for men by 11%.

19. While trends based on age were not consistent, older defendants tended to receive longer sentences overall and juveniles under 17, who had to be waived into the jurisdiction of adult court, tended to receive longer sentences than youth who were just 17.

A. In three offense groups, the largest number of sentences by far was imposed on people who were between 20 and 29. In the CSC group the most common age was 30–39.
B. The median sentence for murder was 20 years for every age except juveniles younger than 17, whose median sentence was three years shorter.
C. Among assault sentences, the median was lowest for those younger than 17 and older than 39.

RECEIVED by MSC 4/12/2022 1:07:49 PM

D. In the CSC and armed robbery groups, juveniles under 17 had longer median sentences than 17-year olds.
    1. Those under 17 were more likely to score at PRV Levels C and D, indicating that the younger defendants had more severe juvenile histories.
    2. For the A grid crimes, those under 17 were more likely to score at OV Levels V and VI than either those who were 17, or the group as a whole, suggesting that the younger defendants committed their offenses in a more serious manner.
    3. Both the youngest groups were more likely to plead guilty than defendants overall.
E. There was virtually no difference by age in the rate of departures.

20. Although non-white incarceration rates are greatly disproportionate to white rates, the causes are complex and rarely appear to result from decisions based directly on race at the sentencing stage.

A. The disparity between white and non-white sentences is most extreme for life-maximum offenses. While non-whites comprise 14% of the state's total population and 55% of the prisoner population, they account for 74% of the life-maximum sentences. This includes: 76% of the murder sentences, 78% of the assault sentences, 79% of the robbery sentences but only 44% of the CSC sentences.
B. Of 2,300 sentences imposed from 1999 to 2012 on defendants who were 17 years or younger when they committed their crimes, 84% involved teenagers of color.
C. Racial statistics vary greatly by county.
    1. In most of the study counties, the population was 18–26% non-white. Wayne County was a key exception at 45.4%.
    2. Five urban counties accounted for 88% of the non-white sentences, with nearly two-thirds coming from Wayne County alone.
    3. In every county the amount of non-white sentences was greatly disproportionate to the county's non-white population.
    4. In most counties the non-white proportion of sentences was greater for each offense type except CSC.
D. There were stark differences between racial groups by age.
    1. Non-whites committed their offenses at a younger age than whites. For example, more than 38% of all the robbery sentences were imposed on non-white defendants who were younger than 20, compared to 6.3% of white youth.
    2. The proportion of non-white sentences was much smaller for CSC than for other offense types at every age. For example, 23.4% of the CSC sentences were imposed on non-whites age 30 and over while 37.1% were imposed on whites that age.
E. Under both the judicial and legislative guidelines, median sentence lengths were significantly longer for non-whites for every life-maximum offense except for murder. The extent of the disparity decreased under the legislative guidelines for CSC and robbery but increased for assault.
F. The extent of significant racial disparity varied by county and offense.
    1. Six counties had no significant instances of disparity.
    2. Five counties had significant disparity for one offense.
    3. Oakland had longer non-white sentences for both assault and robbery.
    4. Wayne imposed significantly longer sentences on non-whites for every offense but murder.

RECEIVED by MSC 4/12/2022 1:07:49 PM

G. Of 10 instances of significant racial disparity, nine involved substantially longer sentences for non-whites. The exception was Genesee County, where the median sentence for robbery was seven years for non-whites and nine years for whites.
H. Racial disparities in sentence length can be explained, at least in part, by significant differences in how routine factors impact each group.
    1. For three of four offense types, non-white sentences fell into higher prior record levels than white sentences. However, non-whites showed less serious prior records than whites for robbery.
    2. Non-whites had significantly higher prior record levels in both Wayne and Oakland Counties.
    3. Overall, higher non-white prior record levels appeared to be primarily attributable to higher non-white scores on PRVs 1, 2, 4, and 6 (high and low severity prior felony convictions, low severity juvenile adjudications, and relationship to criminal justice system).
    4. There were also significant differences in offense severity levels. Whites had consistently lower OV levels for assault and robbery.
    5. Whites had a significantly higher proportion of sentences imposed after guilty pleas than non-whites.
I. Once prior record and offense levels determined the appropriate range, racial disparities could arise from where within, below, or above that range the judge selects the sentence.
    1. Place on the range showed no significant differences by race for most counties.
    2. Statistically significant differences existed for four: Genesee ($p < .01$), Kent ($p < .05$), Oakland ($p < .001$), Saginaw ($p < .001$).
        a. In each of these four counties, non-white sentences were more likely to be both below and at the minimum-minimum than white sentences.
        b. In three of these counties, white sentences were more likely to be at or above the maximum-minimum. In Saginaw, white sentences were more likely to be above the max-min.

21. Michigan's guidelines are designed to be harsher and less consistent than those of other states using guidelines grids.

A. Michigan's lack of a sentencing commission with dedicated professional staff means there is no central body to:

- Monitor the guidelines' implementation.
- Collect, analyze, and publish data regarding their effectiveness.
- Recommend changes necessary to promote the guidelines' goals.
- Respond to changes in the law and public policy.
- Avoid the inconsistencies that tend to arise from ad hoc legislative amendments.

B. Other states have been able to avoid making their guidelines advisory instead of mandatory because their judges engage in fact-finding regarding individual aggravating and mitigating circumstances only after the presumptive sentencing range has been determined.
C. Michigan's presumptive sentencing ranges are far wider than those of other jurisdictions, allowing for substantial disparities in recommended sentences.
D. Michigan is unique in the extent to which it broadens the presumptive sentencing range for habitual offenders.

RECEIVED by MSC 4/12/2022 1:07:49 PM

E. Michigan requires the use of "real offense" sentencing that can increase the punishment for the conviction offense based on charges that were reduced or dismissed as part of a negotiated plea and conduct that was not even charged.
F. Michigan's guidelines repeatedly increase the recommended sentence by counting the same facts multiple times.
G. Whether Michigan's actual sentences are longer than those of other guidelines jurisdictions depends on the extent to which judges choose to depart from the presumptive ranges their guidelines recommend.
    1. However, the construction of Michigan's guidelines was guided by a series of policy choices that consistently favor harsher penalties.
    2. While they often follow the federal example, Michigan's guidelines frequently differ from the more moderate policies of other states.

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER FIVE. HOW SENTENCES GREW LONGER AND MORE NUMEROUS

As we discussed in Chapter One, the legislative guidelines were intentionally designed to increase sentence lengths for life-maximum offenses. In this chapter, we will look at just how much sentence lengths actually grew.

Table 3 shows how each version of the guidelines changed minimum sentences for the life-maximum offenses. Although a major purpose of the second judicial guidelines was to raise the recommended sentences for these crimes, a substantial increase occurred in the median sentence only for murder. The median sentences for CSC and armed robbery were unchanged, and the median for assault actually declined by two years.

From the second judicial to the statutory guidelines, the median sentence increased substantially for every offense group, by a low of 15 months for CSC cases to a high of three years in assault with intent to murder cases. The increases were statistically significant for every offense group except CSC. The percentage increases were greatest for assault and robbery.

Table 3 offers two conclusions. First, compared to the initial guidelines, those enacted by the Legislature 15 years later reflect an increase in sentence lengths of anywhere from 8 to 40%. Second, although the judicial guidelines were revised in the late 1980s specifically to better align sentences for assaultive offenses with the practices of sentencing judges, the sentencing commission that developed the legislative guidelines chose to increase those sentences by 20%.

Table 3. Difference in Median Minimum Sentence between Judicial and Legislative Guidelines, by Offense Type

	M2	AWIM	CSC	RA	Total
Judicial 1 1984–1988	15	12	10	5	7.5
Judicial 2 1989–1998	18	10	10	5	7.5
Difference	**3 (20%)**	**-2 (-17%)*** **	**--------**	**--------*** **	**--------*** **
Legislative 1999–2012	20	13	11.25	7	9
Difference	**2 (11%)*** **	**3 (30%)*** **	**1.25 (12.5%)**	**2.0 (40%)*** **	**1.5 (20%)*** **

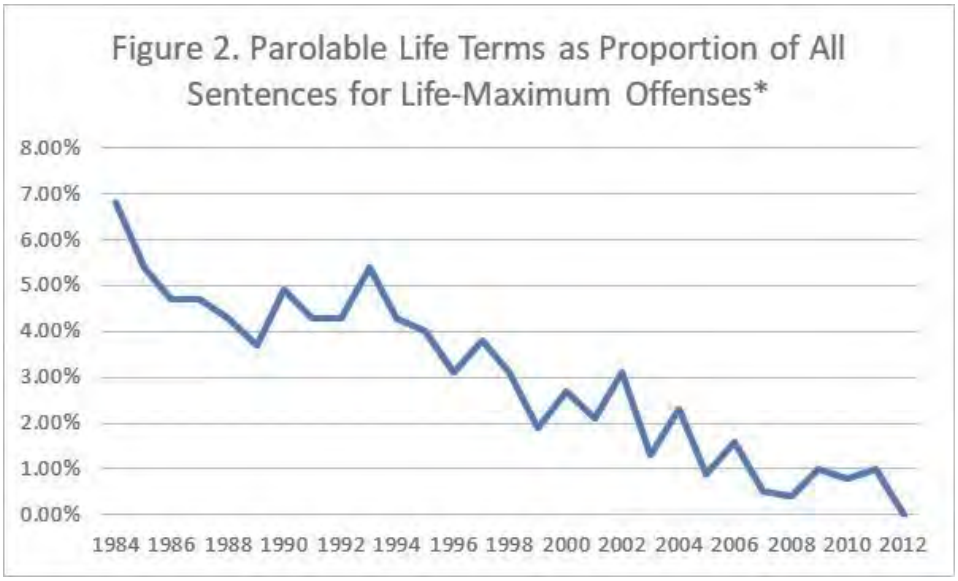
CMIS data, *** $p < .001$

This snapshot provides just the starting point for understanding the impact of the statutory guidelines on sentences for life-maximum crimes. The table masks several important distinctions. First, it does not address the shift from parolable life to long indeterminate sentences. Second, it considers all cases together, regardless of whether the defendant was sentenced as a habitual offender. Third, it fails to recognize large differences among counties. Finally, focusing solely on sentence length fails to capture another phenomenon: the increase over time in multiple sentences per defendant.

A. Life vs. Long Indeterminate Sentences

When we looked in Chapter Two at MDOC commitment data from 1993 to 2017, we saw that sentences of life with the possibility of parole had declined while indeterminate sentences with long minimum terms had increased dramatically. Figure 2 shows that the decline in parolable life sentences began a decade earlier. While life terms were always a small portion of the total sentences actually imposed for life-maximum offenses, the proportion declined from 6.8% in 1984 to 1% or less after 2006. After peaking at 96 in 1993, the annual number of parolable life sentences imposed for offenses committed and sentenced from 2007 to 2012 never exceeded 17.

RECEIVED by MSC 4/12/2022 1:07:49 PM



*CMIS data, includes redundant sentences

The possible reasons for a change in the pattern of life sentences are contradictory and hard to pin down:

- The proportion of cells that allows life as an option actually increased from the second judicial to the legislative guidelines. The difference is greatest by far for second-degree murder; the percentage of cells on the M2 grid that have life as a possibility went from 31% to 72%.
- People sentenced to parolable life terms for offenses committed through October 1992 became eligible for release after serving 10 years. For offenses committed after that date, parole becomes possible after 15 years. Judges who want to ensure that a defendant is incarcerated for a longer period often choose to impose a very long indeterminate term, leading to fewer life sentences and more long minimums over 15 years. Notably, on the M2 grid, the high end of every cell but one substantially exceeds 15 years. On the A grid, 16 of 36 cells (44%) have recommended sentencing ranges with a high end that exceeds 15 years.
- Until the mid-1980s, judges imposed life terms in the belief—based on past parole board practice—that defendants would be released in 12 or 14 years. In fact, lifer paroles had become rare in the late 1980s for various reasons, and in the 1990s, procedures were adopted that deliberately reduced the prospect of a lifer being paroled. Until practices gradually began to loosen again after 2005, virtually no lifers were released except those serving for drug offenses or who were terminally ill. Judges who wanted parole to be a realistic possibility may have begun to impose a term of years with a minimum on which they could rely.[1]

All these scenarios may well be playing out simultaneously. Whatever the reasons, Figure 2 shows that the proportion of life sentences began to decline precipitously after 1993—well before the legislative guidelines were enacted—and continued to drop thereafter.

RECEIVED by MSC 4/12/2022 1:07:49 PM

B. Habitual vs. Non-Habitual Sentences

Recall that the judicial guidelines did not apply to habitual offenders. People were charged and sentenced as habitual offenders during the period that the judicial guidelines were in effect, but the trial judge's selection of a sentence was not made with reference to the guidelines. Conversely, as explained in Chapter One, being sentenced as a habitual offender under the legislative guidelines has enormous consequences. The high end of the applicable guidelines cell is increased, doubling or tripling the breadth of the range for third or fourth offenders and increasing the scope of the judge's discretion accordingly. Lumping habitualized and non-habitualized offenders together produces a marriage of apples and oranges that is misleading for both groups. When we separate the second judicial and statutory periods into non-habitual and habitual sentences in Table 4, four points jump out.

Table 4. Median Length in Years for Two Time Periods – Habitual vs. Non-Habitual Sentences (Number of sentences in parentheses)

	Non-Habitual Sentences Compared by Time Period			Habitual Sentences Compared by Time Period			Habit/Non- Habitual Difference	
Crime	Jud. 2 (Col. A)	Legis. (Col. B)	Change (B/A)	Jud. 2 (Col. C)	Legis. (Col. D)	Change (D/C)	Jud.2 (C/A)	Legis. (D/B)
M2	17.5 (1,891)	20.0 (1,723)	2.5*** 14.3%	25.0 (145)	31.5 (256)	6.5*** 26.0%	7.5*** 42.9%	11.5*** 57.5%
AWIM	10.0 (1,342)	11.25 (1,398)	1.25*** 12.5%	15.0 (188)	25.0 (348)	10.0*** 66.7%	5.0*** 50%	13.75*** 122%
CSC	10.0 (2,314)	10.0 (1,976)	-----***	20.0 (414)	22.6 (454)	2.6*** 13.15%	10.0*** 100%	12.6*** 127.6%
RA	4.0 (7,626)	6.0 (8,883)	2.0*** 50.%	8.0 (1,370)	13.0 (1,913)	5.0*** 62.5%	4.0*** 100%	7.0*** 116.7%
TOTAL	6.5 (13,173)	8.0 (13,980)	1.5*** 23.1%	10.0 (2,117)	15.0 (2,971)	5.0*** 50.0%	3.5*** 53.8%	7.0*** 87.5%
CMIS Data, *** p < .001								

First, for non-habitualized offenders, the increase in the median minimum sentence from the judicial to the legislative period is less substantial for assault than Table 3 suggests and the increase for CSC defendants disappears altogether. However, for all groups combined, the increase in the median sentence for those not charged as habitual offenders was still 1.5 years or 23.0%. The change for every offense group was statistically significant.

By comparison, from the judicial to the legislative period, median sentences for habitualized offenders increased by anywhere from nearly three to 10 years. The percentage increase was greatest for those convicted of assault (66.7%) and robbery (62.5%), whose sentences were the least severe under the judicial guidelines. The change for every offense group was statistically significant. For all offense groups combined, the increase in the median sentence for those charged as habitual offenders was five years or 50.0%.

The sheer fiscal costs of these changes are enormous. In today's dollars, the overall per diem cost per prisoner is $115.09.[2] Over the 14 years during which we have data on the legislative guidelines, 11,434 people not sentenced as habitual offenders had an increase of 1.5 years or 547.75 days in their sentences. This computes to:

RECEIVED by MSC 4/12/2022 1:07:49 PM

11,434 individuals x 547.75 days x $115.09/day = $720,476,635 for 14 years
or $51,462,626 per year.

During the same period, 2,313 people who were sentenced as habitual offenders had an increase of five years or 1,825 days in their sentences. This computes to:

2,313 individuals x 1,825 days x $115.09 = $485,820,785 for 14 years
or $34,701,484 per year.

Thus, the increased habitual and non-habitual offender sentences that resulted from the legislative sentencing guidelines cost taxpayers $86,164,110 per year. That is more than $1.2 billion over 14 years, with no proof that the additional years served increased public safety.

Second, observe how much higher habitual sentences were than non-habitual sentences in each period. When the 2nd judicial guidelines were in effect, habitual offender sentences overall were 54% longer than non-habitual sentences. By offense type, the habitual sentences were 43, 50, 100, and 100% longer than the non-habitual sentences for murder, assault, CSC, and robbery, respectively. Translated into years, this meant that the median cost of being habitualized was 10 additional years for CSC defendants, more than seven years for murder defendants, five years for those convicted of assault, and four years for robbery defendants.

Under the legislative guidelines, the gap between habitual and non-habitual sentences widened much further. For the group overall, habitual sentences were 87.5% longer than those for people not habitualized. In murder cases, the habitual sentences were 57% longer, but in every other offense group the difference was more than 120%. The median cost in years of being habitualized under the legislative guidelines was roughly 13 years for assault and CSC defendants, 11 years for murder defendants, and seven years for those convicted of robbery.

Third, consider the actual minimum number of years being served for each offense group. It is no surprise that murder convictions bring the longest sentences. Both sets of guidelines gave second-degree murder its own sentencing grid with ranges designed to produce longer minimums than other life-maximum offenses. Nonetheless, the median sentence of 31.5 years for those convicted of murder who have been sentenced as habitual offenders is striking, especially since we do not know how old or how serious the prior convictions were.

Also noteworthy are the strong similarities in sentence lengths of those convicted of assault with intent to murder and criminal sexual conduct. When not charged as habitual offenders, the consensus has been consistent that both types of offenses warrant a minimum sentence of 10–11 years. But when charged as habitual offenders, under the legislative guidelines for both groups, the median sentence became nearly 23–25 years.

Equally apparent are the much lower sentences for armed robbery. The consensus is longstanding that armed robbery is a less serious crime than CSC or assault with intent to murder, perhaps because in most armed robberies the victims do not experience actual physical harm.

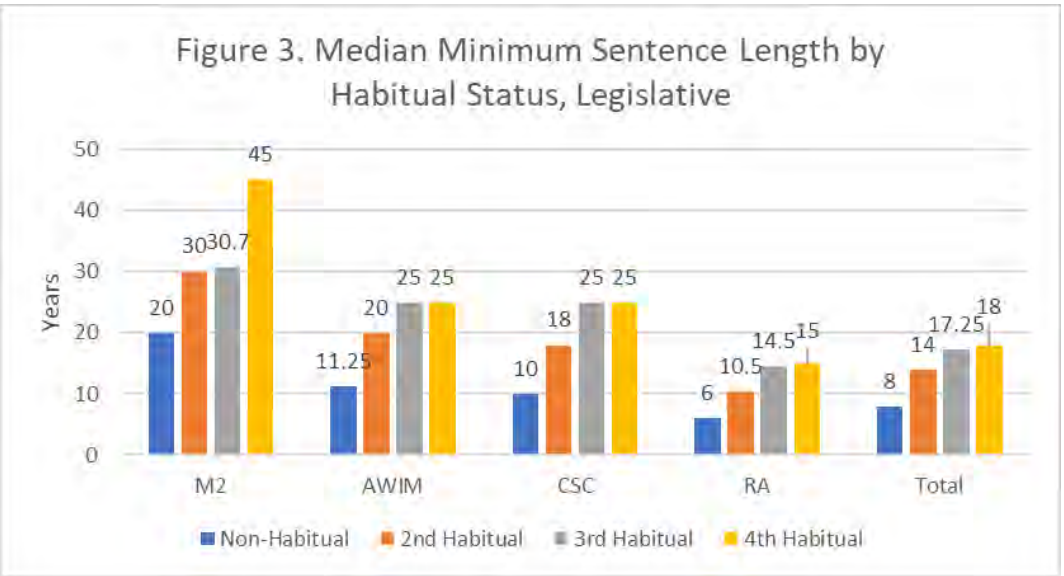
Finally, Table 4 shows a substantial increase in the proportion of life-maximum defendants who are sentenced as habitual offenders. In the 2nd judicial period, 13.8% of life-maximum sentences were imposed with habitual enhancements. In the legislative period, the overall proportion increased by more than a quarter to 17.5%. The overall figure is heavily skewed by the large volume of robbery sentences which had the least growth in habitual charges. By offense, the proportion of habitual offender sentences increased as follows:

RECEIVED by MSC 4/12/2022 1:07:49 PM

- Second-degree murder: 7.1 → 12.9%
- Assault with intent to murder: 12.3 →19.9%
- Criminal sexual conduct: 15.2 → 18.7%
- Armed robbery: 15.2 →17.5%

Both later periods constitute a large increase over the 1980s. When the 1st judicial guidelines were in effect, the proportion of habitual sentences was just 7.4%.

Table 4 compares habitual sentences as a group to non-habitual sentences in different time periods. Figure 3 allows us to see the differences in median sentences in the legislative period depending on whether the defendant was treated as a second, third, or fourth habitual offender.[3]



CMIS Data, All offense types significant at $p < .001$

It is evident that the most significant increases generally occurred as soon as the defendant's status changed from being a non-habitual to a second habitual offender. Murder sentences, which were much longer to begin with, increased by 50%, while minimums in the other offense groups increased by at least 75%. The rate of increase slowed down from second to third habitual. Murder sentences barely changed at all; assault sentences increased by another 25%; CSC and robbery sentences increased by about 38%. From third to fourth habitual—except for murder sentences—the medians changed little or not at all.

These findings have substantial implications for plea negotiations. Although the high end of the guidelines range may increase by 25, 50, or 100%, for people sentenced as second, third, or fourth offenders, the actual medians are 75, 116, and 125% longer than those for non-habtitualized defendants. While a reduction in a habitual charge from fourth offender to third would seem to have great value, except in murder cases the bargain may not mean much. Conversely, a threat to charge someone as a second habitual offender can easily mean a sentence increase of eight years.

C. Differences in Impact Among Counties

Enactment of the legislative guidelines has had very different effects in different counties. Since the judicial guidelines were only advisory, they could be followed to varying degrees by judges around the state. Perhaps even more importantly, the recommended sentence ranges under both sets of guidelines

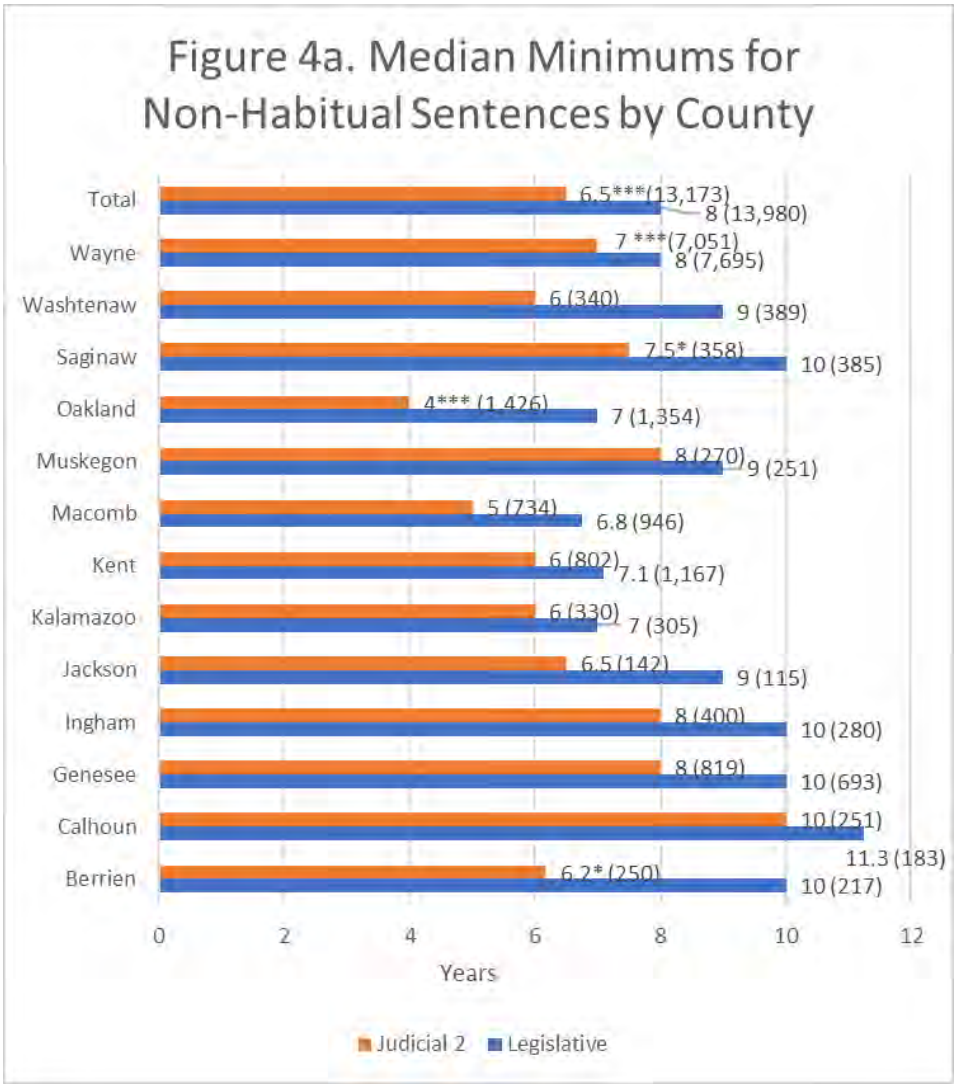
RECEIVED by MSC 4/12/2022 1:07:49 PM

were extremely broad and overlapping. How much the mandatory nature of the legislative guidelines forced a change depended on what local practice had been under the judicial guidelines.

While disparity among counties will be discussed at length in Chapter Ten, Figures 4a and 4b show how the transition from the second judicial to the legislative guidelines affected sentences around the state. This presentation allows us to see which counties saw the most dramatic deviations from then-current practices and which were driving the changes in the statewide averages.

1. Non-Habitualized Defendants

Figure 4a shows that, for those not sentenced as habitual offenders, the median sentence for life-maximum offenses overall increased in every county with the advent of the legislative guidelines. The extent of the increase varied from one year in Wayne, Muskegon, Kalamazoo, and Kent Counties to three years in Oakland and Washtenaw, and 3.8 years in Berrien. Since the counties started in various places, the actual medians remained quite disparate. For example, the Oakland County median went from four to seven years, while Washtenaw County's median grew from six years to nine. The extent of county increases also varied depending on the offense. There are no apparent patterns related to county size or location.



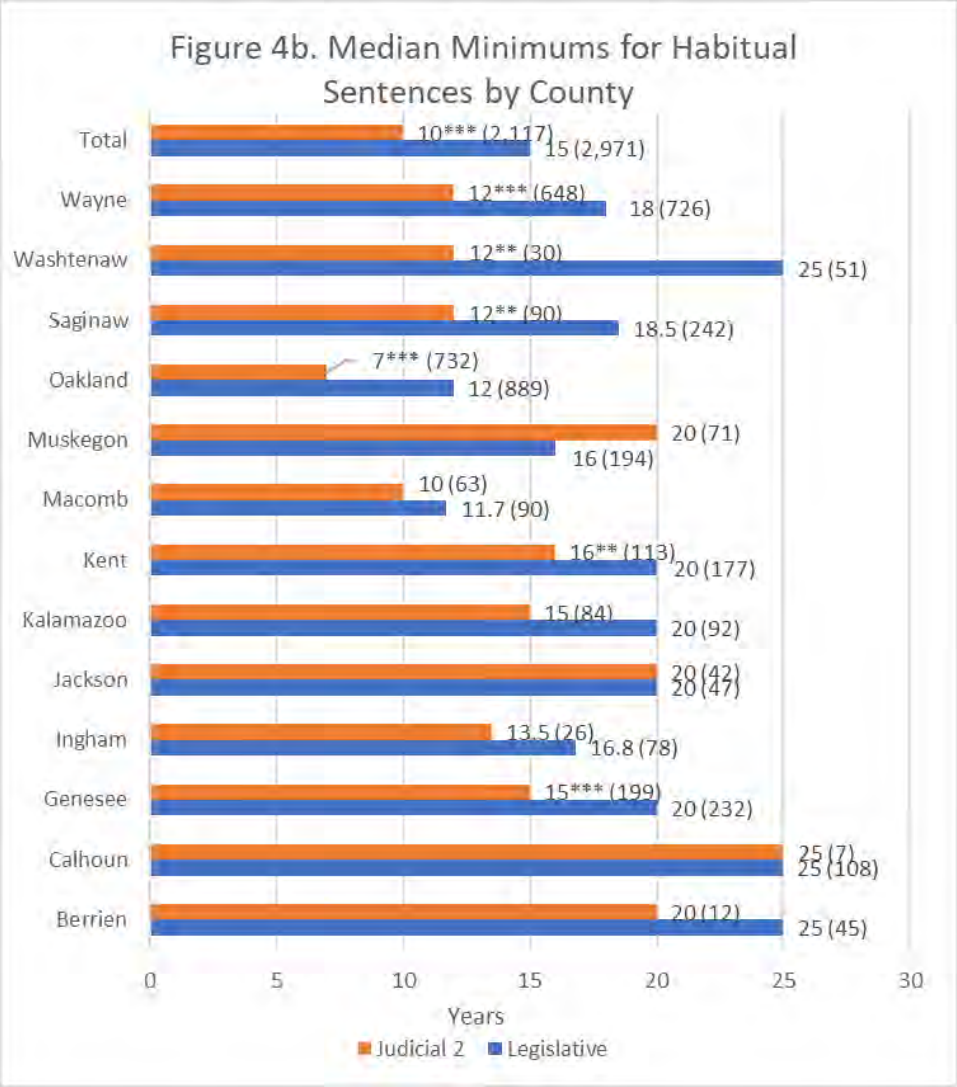
CMIS Data, *p < .05, ***p < .001

RECEIVED by MSC 4/12/2022 1:07:49 PM

2. Habitualized Defendants

Figure 4b shows the change in median sentences across counties for people sentenced as habitual offenders. Given the new application of guidelines to habitual sentences and the seven-year increase in habitual minimums overall, it is not surprising that the changes for most counties were often much greater than they were for non-habituals. The median habitual sentence rose by less than two years in Macomb—from 10 to 11.7 years, while it increased by more than 13 years in Washtenaw—from 12 years to 25. In seven counties, the increase was at least five years. Notably, the median habitual offender sentence declined from 20 to 16 years in Muskegon and stayed constant at 20 years in Jackson and 25 years in Calhoun. There are no evident patterns within counties in the extent to which habitual sentences changed compared to non-habitual sentences.

For those sentenced as habitual offenders, the change in guidelines reduced the differences among the counties by raising the low end of the range. Under the judicial guidelines, the median sentences ranged from a low of seven years in Oakland to a high of 25 in Calhoun—a spread of 18 years. Under the legislative guidelines, the median sentences ranged from a low of 11.7 years in Macomb to a high of 25 in Berrien, Calhoun, and Washtenaw—a spread of 13.3 years.



CMIS Data, **p < .01, ***p < .001

RECEIVED by MSC 4/12/2022 1:07:49 PM

When comparing habitual and non-habitual sentences across counties, there is an important caveat. By definition, defendants sentenced as habitual offenders have prior record levels that reflect some number of felony convictions. Those not sentenced as habitual offenders will have a mixture of prior record levels that may or may not include felony convictions. When counties employ habitual offender enhancements, they pull defendants with the highest PRV levels out of the entire pool. Those left in the non-habitualized group have, on average, lower PRV levels. As a result, counties with a larger proportion of habitual sentences may have relatively low median sentences for non-habituals. Counties that use fewer enhancements may have longer non-habitual sentences precisely because they have not shifted out as many defendants with high PRV levels. This point will be discussed more fully in Chapter Fourteen.

D. The Increase in Multiple Sentences Per Defendant

As explained in Technical Appendix A, for the purpose of determining median sentence lengths, we eliminated multiple sentences that were redundant from the dataset. Nonetheless, examining the frequency of all multiple sentences is important because of the various ways they can be used to make sentences longer.

- Most obvious are the circumstances in which judges have the duty or the discretion to require that sentences imposed at the same time be served consecutively.
- Convictions for multiple charges increase the defendant's prior record score. (See Chapter Six section B)
- Multiple charges and convictions may also increase the offense severity level. (See Chapter Six section C)
- Multiple sentences can also affect the defendant in the future.
    - Each of the multiple sentences can be scored as a prior conviction.
    - Each of the multiple sentences can be used to charge the defendant as a habitual offender.

Thus, even though multiple sentences may be served concurrently, they can provide substantial advantages to the prosecution.

Table 5 displays the total number of sentences and individuals for each guidelines period as well as the frequency of multiple sentences for an individual. Percentages in parentheses show the proportion of individual defendants who received the indicated number of sentences.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 5. Distribution of Multiple Sentences by Habitual Status and Guidelines Period*						
	Judicial 1 (1984–1988)		Judicial 2 (1989–1998)		Legislative (1999–2012)	
Sentences	Non-Habitual	Habitual	Non-Habitual	Habitual	Non-Habitual	Habitual
1	4,149 (83.5%)	355 (84.7%)	8,948 (81.3%)	1,338 (78.6%)	8,459 (74.0%)	1,505 (65.1%)
2	598 (12.0%)	51 (12.2%)	1,464 (13.3%)	231 (13.6%)	1,843 (16.1%)	455 (19.7%)
3	128 (2.6%)	9 (2.2%)	367 (3.3%)	82 (4.8%)	566 (5.0%)	167 (7.2%)
4	52 (1.1%)	3 (0.7%)	121 (1.1%)	17 (1.0%)	268 (2.3%)	75 (3.2%)
5	22 (0.4%)	1 (0.2%)	49 (0.5%)	15 (0.9%)	122 (1.1%)	43 (1.9%)
6	6 (0.1%)	-	22 (0.2%)	9 (0.5%)	56 (0.5%)	26 (1.1%)
7	4 (0.1%)	-	16 (0.1%)	3 (0.2%)	41 (0.4%)	15 (0.6%)
8	3 (0.1%)	-	11(0.1%)	4 (0.2%)	23 (0.2%)	9 (0.4%)
9	2 (0.0%)	-	4 (0.0%)	1 (0.1%)	16 (0.1%)	4 (0.2%)
10 or more	5 (0.1%)	-	7 (0.1%)	2 (0.1%)	40 (0.3%)	14 (0.6%)
Total Sentences	6,228	501	14,154	2,326	16,991	3,974
Total Individuals	**4,969**	**419**	**11,009**	**1,702**	**11,434**	**2,313**

* Note: On this table, individuals are counted more than once if they had sentences that fell into different offense groups.

We can see that over time the practice of charging multiple sentences increased. In the first judicial period, roughly 84% of defendants—whether habitualized or not—had just one sentence for a life-maximum offense. The second judicial guidelines period saw a small decrease in the frequency of single sentences for the non-habitual group and a more marked decline for the habitual group. In the legislative guidelines period, the proportion of defendants who received just one sentence declined to 75% for the non-habitual group and 65% for the habitual group. Conversely, the instances of multiple sentences in every amount grew steadily and substantially.

Further examination reveals that the increased use of multiple sentences differs by offense. Table 6 displays the frequency of multiple sentences for non-habitualized defendants during the legislative guidelines period by offense group. Each group includes all of the sentences imposed on one individual for the indicated type of offense. While 95% of the second-degree murder cases have only one murder sentence imposed, that proportion declines to 84% for assault sentences, 74% for robbery, and 69% for the CSC group. Fifteen percent of the CSC defendants received three or more CSC sentences.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 6. Frequency of Multiple Sentences for Non-Habitualized Defendants by Offense Type under Legislative Guidelines				
Sentence	M2	AWIM	CSC	RA
1	1,629 (95.5%)	1,156 (84.3%)	1,255 (68.9%)	5,268 (74.1%)
2	57 (3.3%)	146 (10.6%)	291 (16.0%)	1,209 (17.0%)
3	17 (1.0%)	36 (2.6%)	137 (7.5%)	299 (4.2%)
4	3 (0.2%)	21 (1.5%)	58 (3.2%)	172 (2.4%)
5	-	5 (0.4%)	27 (1.5%)	63 (0.9%)
6	-	3 (0.2%)	19 (1.0%)	32 (0.5%)
7	-	3 (0.2%)	12 (0.7%)	19 (0.3%)
8	-	1 (0.1%)	5 (0.3%)	15 (0.2%)
9	-	-	8 (0.4%)	7 (0.1%)
10 or more	-	1 (0.1%)	10 (0.5%)	24 (0.3%)
Total Sentences	1,806	1,726	3,061	10,398
Total Individuals	**1,706**	**1,372**	**1,822**	**7,108**

Not surprisingly, multiple sentences appear to be related to conviction method. Except in murder cases, people who pleaded guilty were much more likely to have just one sentence than were those who opted for a jury trial. This difference presumably reflects that reducing the number of charges is one form of plea bargaining. We found no evidence that the number of multiple sentences an individual received was related to race.

Particularly interesting is the extent to which the number of multiple sentences varies by county. As Table 7 shows, the number of defendants who received just a single sentence for robbery ranged from over 90% in Berrien, Jackson, and Kalamazoo, to roughly 67% in Oakland and Macomb all the way down to 53.7% in Saginaw.

For the CSC group, the disparities were much starker. Five counties (Berrien, Genesee, Jackson, Kalamazoo, and Kent) show that at least 80% of defendants received a single sentence. Five other counties (Ingham, Macomb, Muskegon, Washtenaw, and Wayne) ranged from 63 to 69%, while in Oakland and Saginaw, just 54% of the CSC defendants received a single sentence. Indeed, in Saginaw more than a third of the CSC defendants had three or more sentences. At 27.7%, Oakland also far exceeded the extent to which other counties imposed three or more sentences.

In some counties the number of cases is relatively small. However, differences in case numbers do not explain the apparent differences in the number of multiple sentences. When we look at counties with similar case volumes, the disparities remain. Compare, for example, CSC cases in Kent and Oakland or Saginaw and Genesee. A CSC defendant in Oakland is nearly 10 times more likely to receive three or more sentences than a CSC defendant in Kent. A CSC defendant in Saginaw is nearly eight times more likely to receive three or more sentences than a CSC defendant in neighboring Genesee.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 7. CSC/RA Number of Sentences per Individual by County (Legislative, Non-Habituals)								
County	Criminal Sexual Conduct				Armed Robbery			
	1	2	3+	Total	1	2	3+	Total
Berrien	35 83.3%	3 7.1%	4 9.5%	42	127 94.1%	7 5.2%	1 0.7%	135
Calhoun	55 75.3%	9 12.3%	9 12.4%	73	54 80.6%	9 13.4%	4 6.0%	67
Genesee	55 83.3%	8 12.1%	3 4.5%	66	289 81.9%	49 13.9%	15 4.3%	353
Ingham	40 67.8%	10 16.9%	9 10.2%	59	132 88.6%	12 8.1%	5 3.4%	149
Jackson	28 96.6%	1 3.4%	-----	29	67 91.8%	5 6.8%	1 1.4%	73
Kalamazoo	50 82%	7 11.5%	4 6.5%	61	160 90.4%	13 7.3%	4 2.3%	177
Kent	188 80.3%	39 16.7%	7 2.9%	234	568 83.5%	90 13.2%	22 3.2%	680
Macomb	85 63.4%	2 19.4%	23 17.2%	134	369 67.8%	120 22.1%	55 10.2%	544
Muskegon	52 69.3%	14 18.7%	9 12.0%	75	100 83.3%	16 13.3%	4 3.1%	120
Oakland	119 53.8%	41 18.6%	61 27.7%	221	487 66.9%	167 22.9%	74 10.2%	728
Saginaw	34 54.0%	7 11.1%	22 35.1%	63	95 53.7%	59 33.3%	23 13.0%	177
Washtenaw	39 68.4%	9 15.8%	9 15.9%	57	158 71.2%	44 19.8%	20 9.2%	222
Wayne	478 67.1%	119 16.7	115 16.1%	712	2,827 73.8%	634 16.5%	372 9.7%	3,833
CMIS Data	p < .001				p < .001			

Variation in the number of sentences imposed is just one example of how the county of conviction affects defendants. In Chapter Ten, we will explore the relationship between county and the median sentence length. In Chapter Fourteen, we will examine disparities in the rates at which counties choose to utilize the habitual offender statutes. The outcomes are sometimes surprising. But the result is always that—statewide sentencing guidelines notwithstanding—the exercise of discretion by local courts and local prosecutors produces widely disparate results.

Conclusion

Implementation of the legislative guidelines had multiple consequences.

- The median length of non-habitual sentences grew overall by 1.5 years.
- The median length of habitual offender sentences grew overall by five years.
- The proportion of habitual offender sentences increased by more than a quarter.
- The imposition of multiple sentences for individual defendants increased dramatically.

The extent of all these changes varied substantially by offense group and by county. But the trends are undeniable. For the life-maximum offenses, the legislative guidelines brought longer sentences, as they were designed to do. In the next chapter, we will examine that design in detail to see exactly how it achieved these results.

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER SIX: HOW THE GUIDELINES PROMOTE LONGER SENTENCES

The longer sentences for life-maximum offenses were achieved by carefully constructing the guidelines components and anticipating the intricate interplay among them. Because there has been no ongoing commission monitoring, the impact of each component has not been evaluated.

Five factors determine the parameters of the recommended sentencing range:

- The grid on which the crime type falls.
- The prior record score, which places the defendant into one of six prior record levels (A–F).
- The offense score, which places the defendant into one of multiple offense severity levels, depending on the grid.
- The cell range that sits at the point on the grid where the variable levels intersect.
- Whether the high end of the cell is increased because the defendant is being sentenced as a habitual offender.

We saw in Table 4 how the number and length of habitual offender sentences exploded when the statutory guidelines were enacted. In Chapter Fourteen, we will examine who is chosen to be habitualized. Here we will deconstruct that design in order to examine exactly how the grids, the prior record variables, and the offense variables each work to promote longer sentences, along with policy decisions about habitual offenders and departures.

A. The Grids

Michigan divides all offenses into six crime groups based on the general nature of the conduct.[1] The crime group is used to determine which offense variables are to be scored. Michigan also assigns offenses to one of nine crime classes, based roughly on the statutory maximum sentence. Each crime class has its own sentencing grid. Among the life-maximum offenses, murder is treated separately and has its own grid, M2. The other life-maximum offenses are on the A grid. Grids B–H, in descending order of seriousness, are for offenses with maximums of 20 years in prison, down to jail or other intermediate sanctions. Thus, the grid is a proxy for the relative seriousness of the offense type.

1. Sentence Increases from Grid to Grid

The recommended minimum sentences change dramatically from grid to grid. Consider, for example, a defendant whose PRV Level is C. (Note that the B grid was designed for offenses with statutory maximums of 20 years.)

- If the defendant scores 50 OV points, the minimum-minimum goes from 4.25 years on the B grid to 6.75 on the A grid to 15 years on the M2 grid.
- If the defendant scores 100 OV points, the minimum-minimum on the B grid increases by just six months to 4.75 years, but it goes to 11.25 years on the A grid and 18.75 years for murder 2.

Similarly, consider a defendant whose PRV Level is D.

- If the defendant scores 50 OV points, the minimum-minimum goes from 6.5 years on the B grid to nine on the A grid to 18.75 years on the M2 grid.

RECEIVED by MSC 4/12/2022 1:07:49 PM

- If the defendant scores 100 OV points, the minimum-minimum on the B grid again increases by just six months to seven years, but it goes to 14.25 years on the A grid and 22.5 years for murder 2.

The percentage change in the range from grid to grid varies substantially.

- Moving from the B to the A grid, the greatest changes occur when offense severity increases.
- Moving from the A to the M2 grid, the greatest changes occur when PRV level increases.

In every cell, the maximum-minimum is 66.7% greater than the minimum-minimum. The low end of each range is 25% below the midpoint, and the high end is 25% above the midpoint. As the midpoints grow larger, the actual increases in the widths of the ranges are stunning.

- On the A grid, the range for the A-I cell is 14 months; for the F-VI cell it is 15 years.
- On the M2 grid, the narrowest range is five years and the widest is nearly 20.

2. The Number and Redundancy of Cells

As explained in Chapter One, all the grids have six PRV levels. The M2, G, and H grids have three OV levels, for a total of 18 cells each. The A, B, C, D, and E grids have six OV levels, for a total of 36 cells each. The F grid has four OV levels, for a total of 24 cells. Thus, collectively, the Michigan guidelines present a choice among 258 cells.[2]

The grids are structured so that PRV levels A–F are arranged horizontally across the top of the grid, while the OV levels run vertically down the left side. On grids A–D, the cell ranges in each row have a column that is an exact counterpart. For example, on the A grid, the low ends of the cells in the row for OV Level III are 42, 51, 81, 108, 126, and 135 months. The low ends of the cells in the column for PRV Level C are also 42, 51, 81, 108, 126, and 135 months. The pattern is consistent throughout the grid. The column of cells for PRV Level A matches the row of cells for OV Level I; the column for PRV Level F matches the row for OV Level VI. The same concept is employed on the M2 grid but adapted to account for the fact that the grid has only three OV levels. To visualize this fully, see Figures 5a and 5b in Chapter Eight.

This repetition produces a series of diagonals along which different combinations of PRV and OV levels result in identical cells. Thus cells A-VI, F-I and every combination along the diagonal line between them has a range with a low of 108 months and a standard high of 210 months. As a result of this design, the M2 grid has 18 total cells but only eight different recommended minimum sentence ranges. On the A grid, of 36 total cells, there are actually just 11 different ranges.

The apparent intent of all this redundancy is to treat prior record and offense severity as equally important. A balance is repeatedly struck between cases that have more prior record points but fewer OV points and cases that have fewer prior record points but more OV points. Chapter Eight explores how successful this strategy has been at producing equivalent median sentences. But, from a broader perspective, there is reason to question the desirability of a policy that makes the seriousness of the current offense no more important than the defendant's criminal history, no matter what combination of factors produced the prior record score. A person with a long history of relatively petty crimes may receive a longer recommended sentence than someone with a more serious current offense. Longer prison sentences can be based on prior record scores that reflect a mix of factors over a long period of time but not a single assaultive felony. It is worth reconsidering whether this is the best use of scarce resources, the best way to protect the public, or the fairest way to treat defendants.

The guidelines ranges do not just repeat, they overlap. Adjacent cells do not begin where the next lowest one left off. Instead the very wide ranges include many of the same months. For instance, the C-III

RECEIVED by MSC 4/12/2022 1:07:49 PM

cell on the A grid is 81–135 months; the C-V and D-III cells are 108–180 months. The period of 108–135 months is included in both ranges—an overlap of 27 months.

The confounding impact of overlapping and redundant cells is apparent if we consider a minimum sentence of 180 months or 15 years. It is included in the ranges starting with 108, 126, 135, and 171 months. Collectively, that is 18 of the A grid's 36 cells. What, exactly, is a 15-year minimum meant to punish when it is equally available at one cell at OV Level I, two cells at OV Level II, three of cells at OV Level III, and four cells each at OV Levels IV, V, and VI?

B. Prior Record Variables

Before examining the nature and impact of each of the seven prior record variables, it is helpful to consider some of the important policy analysis contained in the Robina Institute's "Criminal History Enhancements Sourcebook."[3] There are two primary rationales for increasing a sentence based on the defendant's criminal history. One is retribution: the notion that people who continue to commit crimes despite previous encounters with the criminal justice system are more blameworthy. The other is risk reduction: the notion that longer sentences reduce the risk of re-offending.[4]

Sentencing commissions are typically not explicit about which rationale underlies their choices. Presumably it is often both. But the rationales lead to different policy choices. If the goal is risk reduction, it makes sense to assess the effectiveness of each criminal history component at actually reducing future offending. Validation research based on data available in each jurisdiction can be used to determine whether and to what extent a prior record variable affects re-offending.[5]

Such research is critical because criminal history enhancements can have a number of adverse impacts.[6]They can:

- Increase the size and expense of prison populations.
- Shift the age composition and risk level of prison inmates as older, non-violent individuals tend to have longer criminal histories.
- Undercut the goal of using limited prison beds for people convicted of violent crimes.
- Decrease the proportionality of sentence severity relative to offense severity; that is, the more weight given to prior record, the less the sentence length depends on the severity of the conviction offense.
- Increase racial disparity in prison populations because, as will be discussed in Chapter Thirteen, non-white defendants have larger criminal history scores than whites.

To assess the impact of Michigan's prior record variables, we will refer to Table 8, which shows the extent to which no points were scored for each PRV for sentences in each offense group.

Table 8. Percent of Non-Habitual Sentences with Zero Points for Each Prior Record Variable under Legislative Guidelines, 2003–2012

Prior Record Variable		Murder 2nd	Assault	CSC 1st	Robbery
1	High Severity Felonies	81.7	77.7	85.1	81.4
2	Low Severity Felonies	59.9	56.6	76.9	64.3
3	High Severity Juvenile	93.3	92.0	94.8	89.6
4	Low Severity Juvenile	80.1	78.6	93.2	76.8
5	Misdemeanors	54.7	54.2	63.4	53.6
6	Relationship to Criminal Justice System	64.4	60.9	81.6	60.6
7	Subsequent or Concurrent Felonies	65.0	46.9	43.5	52.7
OMNI data, $p < .001$					

RECEIVED by MSC 4/12/2022 1:07:49 PM

Prior adult felonies. Every version of the guidelines scored both high and low severity adult felonies. At each severity level, scoring is the same regardless of whether the prior conviction resulted in probation, jail, or prison; that is, the fact of the prior conviction not the seriousness of the specific conduct determines the points.

The statutory guidelines define high severity felonies as those listed in crime classes M2, A, B, C, and D (i.e., crimes carrying maximum sentences of over five years). According to the MDOC 2019 Statistical Report,[7] the most common high severity crimes are possession of methamphetamine,[8] delivery of narcotics under 50 grams,[9] assault with intent to commit great bodily harm less than murder,[10] breaking and entering a commercial structure,[11] and uttering and publishing.[12]

There is no requirement that prior felonies must have arisen from different criminal incidents; that is, if multiple convictions resulted from a single transaction, such as three counts of armed robbery because there were three victims or car theft, possession of drugs and fleeing from police, each of those convictions is scored as a separate prior. There is also no prohibition on scoring prior felonies that were used to enhance the punishment for the current offense. While this is not an issue with the life-maximum offenses, there are many less serious crimes for which the statutory penalty is increased if the defendant committed the same crime in the past. In that circumstance, the prior conviction is used to enhance the person's punishment at least twice.[13]

PRV 1. The statutory guidelines award 25 points for one prior high severity felony conviction, 50 points for two, and 75 points for three or more. At 25 points, one prior high severity felony alone puts the defendant into Level D, and two such felonies push the defendant into at least Level E. As Table 8 shows, more than three-quarters of sentences in every crime group had no prior high severity felonies, with the CSC group having the fewest and the assault group having the most. At the other extreme, the proportion of sentences that score 75 points ranged from 1.8% in CSC cases to 3.7% in armed robbery cases. From 9.1 to 15.3% received 25 points in CSC and assault cases, respectively; from 3 to 5% received 50 points.

PRV 2. PRV 2 awards five points for a single low severity felony conviction and a maximum of 30 points for four or more. More than half of all sentences received no points for low severity priors but the proportion varied widely by offense group, ranging from a low of 56.6% in assault cases to 76.9% in CSC cases. The proportion receiving either five or 10 points was 16.2% for the CSC group but ranged from 25.9 to 30.9% for the other groups. Except for CSC sentences, 10.8–12.6% of sentences were awarded 20 or 30 points, enough to put them into PRV levels C or D, respectively.

Juvenile adjudications. The legislative guidelines divide juvenile adjudications into high and low severity on the same basis as adult convictions. However, the point scale for juvenile adjudications is substantially lower than for adult convictions in recognition of juveniles' lesser culpability. Juvenile adjudications do not "roll off" at any point simply because they were juvenile. They are subject to the same 10-year gap rule as adult convictions.

PRV 3. The guidelines award 10 points for a single high severity juvenile adjudication, 25 points for two and 50 points for three or more. Table 8 shows that roughly 90 to 95% of the sentences in every offense group had no points for this variable. Fewer than 3% in any group scored 25 points or more.

PRV 4. The guidelines award two points for a single low severity juvenile adjudication up to a maximum of 20 points for six or more. Table 8 shows that over 93% of CSC sentences and about 77 to 80% of the sentences in every other offense group had no points scored for this variable. Except for sex offenses, which scored far fewer, roughly 12-13% of sentences were scored two points for one low severity juvenile adjudication. Just 3.25% of all sentences received either 15 or 20 points.

Misdemeanor convictions (PRV 5). The statutory guidelines award two points for one prior misdemeanor and up to 20 points for seven or more. Just three misdemeanors in the preceding 27 years brings 10 points—enough, in itself, to put the defendant at Level C. However, nearly 87% of all the life-maximum sentences scored no more than five points for misdemeanors. Roughly 54% of the sentences in three of the offense groups scored zero; among CSC sentences, the proportion was 63.4%.

The statutory guidelines score any misdemeanor or equivalent juvenile adjudication that is an offense against a person or property, a controlled substance or weapons offense, or involves operating a vehicle of any kind while under the influence. No differentiation is made based on the severity of the

RECEIVED by MSC 4/12/2022 1:07:49 PM

conduct. However, a prior misdemeanor that was used to elevate the sentencing offense to a felony, as occurs, for instance, with first-degree retail fraud and third-offense drunk driving cannot be scored.

Eight of 17 guidelines jurisdictions that count juvenile adjudications limit the scoring to felonies. Michigan is one of nine jurisdictions that count juvenile misdemeanors.[14] In fact, it awards almost identical points for misdemeanors regardless of the defendant's age when committed and low severity juvenile adjudications: two points for one of either, five points for two of either, 10 points for three or four of either, and 15 points for five of either. The point scheme only diverges at the high end, where six or more low severity juvenile adjudications are awarded 20 points but seven or more misdemeanors are required for 20 points.

Given that juveniles' reduced culpability is taken into account in scoring juvenile adjudications, it is unclear why they are treated like adults in scoring misdemeanors. Two prior misdemeanors committed as a juvenile (five points) and two prior low severity juvenile adjudications (five points) are enough to push the defendant into Level C. On the A grid, compared to PRV Level A this combination doubles both the low and high ends of the recommended ranges at Offense Levels I-IV. For example, at OV Level III, a minimum of 3.5 years (42 months) becomes 6.75 years (81 months).

Criminal Justice Status (PRV 6). PRV 6 awards points based on the defendant's relationship to the criminal justice system when the sentencing offense was committed. The four categories of points depend on whether the defendant was incarcerated and, to a limited extent, whether the relationship arose from a felony or misdemeanor:

- Five points if the defendant was on probation, delayed sentence status, or out on bond regarding a misdemeanor.
- Ten points if the defendant was on parole, probation, delayed sentence status, or out on bond regarding a felony.
- Fifteen points if the defendant was in jail awaiting adjudication or sentencing on a conviction or probation violation (no difference whether the pending proceeding was for a felony or misdemeanor).
- Twenty points if the defendant was serving a sentence in prison or jail (includes escapees; no difference in whether jail sentence was for a felony or misdemeanor).

Thus, serving a jail sentence is treated as the equivalent of three prior low severity felony convictions. Being on probation for a single low severity felony is the equivalent of two prior low severity felony convictions.

The rationale for PRV 6 is unclear. While a common factor in sentencing guidelines scoring, custody status is the least researched. It is not considered a predictor of re-offending in most risk assessment instruments.[15] Its one clear result is to pile punishment on top of punishment.

Every relationship to the criminal justice system already carries its own consequences for the defendant who has committed a new felony. Probation, parole, and bond can be revoked; sentences can be made consecutive. MCL 768.7a requires that a sentence imposed for an offense committed while the defendant was incarcerated or on parole be served consecutively to the prior sentence. Thus, for instance, a defendant who is sentenced for being a prisoner in possession of a cell phone,[16] which carries a maximum sentence of five years, will be scored 20 points for PRV 6, 25 points for OV 19 (Threat to Security of a Penal Institution), and the sentence will be consecutive to the one being served.

As Table 8 shows, nearly 82% of the CSC sentences and more than 60% of all others were awarded zero points for PRV 6. In every offense group, 1.3% or less of the sentences received 15 or 20 points combined; however, except for the CSC group, between 25 and 31% of sentences received 10 points, putting them immediately into PRV Level C.

Subsequent/Concurrent Convictions (PRV 7). While PRVs 1 and 2 are careful to define "prior" convictions as ones entered before the sentencing offense was committed, PRV 7, called "Subsequent or Concurrent Felony Convictions," undercuts that concept. PRV 7 awards points for two types of

RECEIVED by MSC 4/12/2022 1:07:49 PM

convictions that, by definition, are not prior: if the defendant was convicted of multiple felony counts including the sentencing offense or was convicted of a felony after the sentencing offense was committed. Ten points are awarded for one concurrent or subsequent conviction; 20 points are awarded for two or more. Convictions for felony firearm or for other offenses that result in mandatory consecutive sentences are not counted. There is no exception for convictions that result in discretionary consecutive sentences.

PRV 7 is the single most impactful PRV as a practical matter. Table 8 shows that while 65% of the murder sentences scored zero points for PRV 7, the proportions for the other offense groups were much lower. In fact, for the other three groups PRV 7 had a lower proportion of zero scores than any of the other PRVs. The contrast was starkest for CSC sentences, which had the greatest percentage of zero scores on every other PRV and, at just 43.5%, the lowest percentage of zero scores on PRV 7. PRV 7 has the greatest impact on CSC sentences because prosecutors commonly charge separate counts for multiple penetrations of the same victim in a single incident or of the same or multiple victims over time.

Overall, 21.4% of the sentences under the statutory guidelines were awarded 10 points, and 26% were awarded 20 points, meaning 47% were pushed into at least PRV Level C based on this single variable.

PRV 7 has two purposes. It nullifies some of the benefits of concurrent sentencing. Although the defendant who has multiple convictions out of a single transaction may only have to serve the longest sentence imposed, by increasing the number of points used to calculate that sentence, the other convictions are used to enhance the punishment. Unlike the PRVs that actually focus on the defendant's past, PRV 7 gives substantial control to prosecutors who decide how many concurrent charges to bring.[17]

Similarly, subsequent convictions cannot be scored as prior convictions and cannot be the basis for charging someone as a habitual offender. But PRV 7 allows them to be scored to enhance the defendant's punishment nonetheless. Of course, when the defendant is sentenced for the subsequent conviction, the current conviction will be scored as a prior conviction for purposes of enhancing the subsequent sentence.

As will be discussed in more detail below, PRV 7 can also result in the double-counting of convictions as both prior record and offense variables. Concurrent and subsequent convictions that are scored in PRV 7 may also be scored as part of a continuing pattern of criminal behavior under OV 13. In sum:

- A great majority of life-maximum sentences receive zero points for high severity felonies or high or low severity juvenile adjudications.
- A majority of sentences in every crime group have no low severity felonies, no misdemeanors, and no relationship to the criminal justice system.
- On six prior record variables, CSC sentences are far and away the most likely to receive zero points.
- PRV 7, which adds 10 or 20 points for concurrent or subsequent convictions that are, by definition, not "prior" to the instant offense is the single most consequential.

Of course, what ultimately counts is the total number of points a defendant scores on all the PRVs—that determines in which of the six prior record levels the case is placed. Table 9, which indicates the number of points that define each variable level in parentheses, shows that under the statutory guidelines, for people not sentenced as habitual offenders, the single most common PRV level for three offense groups is C; for the assault group, it is D. Overall, CSC defendants have by far the lowest prior record scores and assault defendants have the highest. While at least 50% of the sentences for every offense type fell at Level C or below, this ranged from just 50.5% for assault sentences to 70.9% for CSC sentences.

Roughly 15–20% of the sentences fall into Level A, although that requires the defendant receive no points for any prior record variable. When we collapse Levels A and B, we see the extent to which sentences in each group have nine or fewer points:

RECEIVED by MSC 4/12/2022 1:07:49 PM

- Murder, 2nd degree – 33.2%
- Assault with intent to murder – 23.0%
- Criminal Sexual Conduct, 1st degree – 27.3%
- Armed Robbery – 25.7%

Table 9. Percentage of Non-Habitual Sentences Among PRV Levels, by Offense 2003–2012

PRV Level (points)	M2 (N = 1,218)	AWIM (N = 884)	CSC (N = 1,225)	RA (N = 5,659)	TOTAL (N = 8,986)
A (0)	19.8	14.9	20.0	15.3	16.5
B (1–9)	13.4	8.1	7.3	10.4	10.1
C (10–24)	29.1	27.5	43.6	31.4	32.3
D (25–49)	22.2	28.5	16.2	24.0	23.1
E (50–74)	9.1	11.9	7.8	10.8	10.3
F (75 +)	6.5	9.0	5.1	8.1	7.6
OMNI Data, p < .001					

At the other extreme, only 5 to 9% fall into Level F. More than 82% of the nearly nine thousand non-habitual sentences for which guidelines scores are available did not reach PRV Levels E or F.

The statutory guidelines were designed to place defendants with one prior high severity conviction into at least PRV Level D and to place defendants with one prior high severity juvenile adjudication or one prior or subsequent conviction of any felony or nearly every possible current relationship to the criminal justice system into at least PRV Level C. It does not appear that, overall, the prior record variables are increasing sentence lengths by driving defendants into the highest guidelines ranges. The majority of life-maximum defendants have criminal histories that are minimal or relatively modest.

Nonetheless, since prior record level depends on the total number of points, regardless of how they accumulated, it is quite possible for a defendant to score into PRV Level E without any prior convictions more serious than a Class E felony with a five-year maximum.[18] For the 41% of sentences that score at PRV Level D or above and particularly for the 18% that score in the E and F levels, the scoring of the prior record variables can serve to push defendants with relatively low offense severity scores into high guidelines cells, adding years to the low end of the recommended range.

C. Offense Variables

The Michigan Felony Sentencing Project found that, for serious assaultive offenses, the facts of the crime play a much larger role in determining the sentence length than does the defendant's criminal history.[19] The results of our analysis of the legislative guidelines are less clear cut.

The 20 offense variables used to score crimes against people address every possible aspect of the defendant's conduct. Some apply to person crimes generally but have exceptions for specific offenses. Two are scored only if the offense involved the operation of a vehicle or other means of transportation. Four carry a maximum of 15 points; five carry a maximum of 10 points. At the other extreme, four OVs carry up to 50 points and two carry up to 100.

The offense variables used to score life-maximum crimes evolved with each version of the guidelines. Most notably, the statutory guidelines added four entirely new offense variables:

RECEIVED by MSC 4/12/2022 1:07:49 PM

- OV2 Lethal Potential of Weapon Possessed or Used
- OV 13 Continuing Pattern of Criminal Behavior
- OV 19 Threat to Security of Penal Institution or Court or Interference with Administration of Justice or Rendering of Emergency Services
- OV 20 Terrorism

The Guidelines Manual conveys the Commission's intent to maximize the number of OV points awarded to every defendant. The instructions direct: "Determine which one or more of the statements addressed by the OV apply to the offender and assign the point value indicated by the applicable statement with the highest number of points."[20]
Table 10 shows that 53% of murder defendants and nearly 45% of assault defendants receive 100 or more points, thereby falling into the highest offense levels on their respective grids. Within the sex offense group, nearly 46% fall into the two highest offense levels on the A grid. For these three offenses, the lowest offense levels are nearly irrelevant. Less than 3% of the CSC sentences fall in OV Level I; less than 5% of the assault sentences fall in either OV I or OV II; less than 7% of the murder sentences fall into OV I which, on the M2 grid, includes up to 49 points. Robbery cases, however, present the opposite scenario with nearly half falling in the two lowest OV levels and more than three-quarters falling in the lowest three.

Table 10. Percentage of Non-Habitual Sentences Among OV Levels, by Offense 2003–2012

OV Level (points)	M2 (N = 1,218)	AWIM (N = 884)	CSC (N = 1,225)	RA (N = 5,659)	Total
I (0–19)		1.9	2.5	11.9	8.9
M2 (0–49)	6.7*				
II (20–39)		2.9	12.7	36.2	30.2
M2 (50–90)	39.8*				
III(40–59)		9.4	16.7	27.8	28.0
M2 (100+)	53.4*				
IV (60–79)		20.8	22.3	12.8	13.2
V (80–99)		20.4	16.5	5.0	7.4
VI (100+)		44.6	29.2	6.2	12.3
OMNI Data, p<.001					
*Note that the M2 grid has only three offense levels with point totals divided differently than the A grid					

These figures, while striking, are not a surprise. For each crime type, the OVs were structured to produce just these results. To examine the impact of offense variables on sentences for life-maximum crimes, we will highlight aspects of the 18 OVs that apply to our four life-maximum offenses.

Aggravated Use of Weapon (OV 1). OV 1 scores 25 points for discharging a firearm at a victim or for cutting or stabbing a victim. Nearly 75% of murder sentences and 80% of assault sentences had 25 points. Over 61% of robbery sentences had 15 points for pointing a gun at or toward a victim or threatening a victim with a knife. On the other hand, 88% of the CSC sentences had zero points for this OV. OV 1 also allows 10 points if the victim was touched by any other type of weapon and five points if a weapon was displayed or implied. Accomplices who were unarmed receive the same number of points as the principal.

RECEIVED by MSC 4/12/2022 1:07:49 PM

OV 1 is scored if the sentencing offense is armed robbery or felonious assault, both of which have possession or use of a weapon as an element, and if the defendant is also receiving a mandatory consecutive sentence for possession of a firearm during the commission of a felony (felony firearm). [21] Notably, in defining permissible sentencing considerations, the *Model Penal Code: Sentencing* would require that "no finding of fact may be used more than once."[22]

Lethal Potential of Weapon Possessed or Used (OV 2). OV 2 goes beyond scoring the use of a weapon done in OV1 to add points depending on the nature of the weapon:

- Fifteen points for possessing chemical or biological weapons, incendiary or explosive devices, and fully automatic weapons,
- Ten points for short-barreled rifles or shotguns,
- Five points for pistols, rifles, shotguns, and knives,
- One point for any other potentially lethal weapon

Thus, a defendant who pointed a handgun at a victim during a robbery would be sentenced on the A grid for being armed during the robbery, then receive 15 points for pointing the gun under OV 1, and get an additional five points under OV2, immediately pushing his sentence into OV Level II. Among murder, assault and robbery sentences, 70% or more received five points for OV 2. (Conversely, nearly 90% of CSC sentences received zero points.)

As with OV 1, all co-defendants must be assessed the same number of points for possessing a weapon, regardless of whether they personally were armed. Also as with OV 1, a defendant who is being sentenced for armed robbery or felonious assault can receive points for OV 2 and also receive a mandatory consecutive two-year sentence for felony firearm.

Physical Injury to Victim (OV 3). OV 3 scores 100 points if death results from commission of a crime and homicide is not the sentencing offense. So, for instance, if the defendant is an accomplice to a robbery or assault in which the principal kills someone, and the defendant is allowed to plead guilty to robbery or assault, he will still receive 100 points for the death. One hundred points were scored in 7.6% of assault sentences and 1.9% of robbery sentences. Additional scoring under OV 3 is:

- Fifty points if death resulted from the operation of any sort of vehicle when the defendant was under the influence of drugs or alcohol.
- Twenty-five points for life threatening or permanent incapacitating injury. Twenty-five points can also be scored when the sentencing offense is homicide and the defendant causes physical injury to the victim in the process of the killing.
- Ten points if the victim incurred bodily injury requiring medical treatment.
- Five points for bodily injury that did not require treatment. Notably, five points cannot be scored if bodily injury is an element of the offense but 10 points can be.

Nearly 30% of murder sentences scored 50 points; 23.2% scored 25. Among assault sentences, nearly 40% scored 25 points; nearly 34% scored 10. Nearly two-thirds of CSC sentences and three-quarters of robbery sentences scored zero on OV 3.

Psychological Injury to Victim (OV4). OV 4 awards 10 points if "serious psychological injury requiring professional treatment occurred to a victim." "Serious psychological injury" is not defined. The instructions require these points to be scored if the injury "may" require treatment; they do not require that treatment actually had been sought. OV 4 applies to all offenses except those involving a controlled substance. The percentage of sentences that received 10 points were: murder – 5.6, robbery – 24.7, assault – 40.0, and CSC – 85.3%.

Psychological Injury to Member of Victim's Family (OV 5). OV 5 awards 15 points if "serious psychological injury requiring professional treatment occurred to a victim's family." As with OV 4, "serious psychological injury" is not defined, and treatment need not actually have been sought. OV 5 is

RECEIVED by MSC 4/12/2022 1:07:49 PM

scored only if the sentencing offense is homicide, attempted homicide, conspiracy or solicitation to commit a homicide, or assault with intent to murder. Fifteen points were scored in 16.9% of the assault sentences and 63.0% of the murder sentences.

Offender's Intent to Kill or Injure Another Individual (OV 6). OV 6 is scored only if the offense is homicide, attempted homicide, conspiracy, or solicitation to commit a homicide or assault with intent to commit a homicide. It awards points for the same state of mind that is an element required for conviction of these offenses. Because of the intent to kill or injure the crimes of second-degree murder and assault with intent to murder are placed, respectively, on the M2 and A grids, which recommend harsher sentences. OV 6 then increases the recommended sentence based on that same intent.

Notably, the *Model Penal Code: Sentencing* would prohibit departure sentences "based on any factor necessarily included in the elements of the offense of which the offender has been convicted."[23] However the Michigan Supreme Court has taken the opposite position. Observing that the sentencing guidelines explicitly direct courts not to consider conduct inherent in a crime when scoring certain offense variables, the Supreme Court held that "absent an express prohibition, courts may consider conduct inherent in a crime when scoring offense variables."[24]

OV 6 scores 50 points for premeditation, or if the killing occurred during an enumerated felony. Thus 50 points can be scored when first-degree murder is pleaded down to second-degree and when the sentencing offense is assault with intent to murder. It scores 25 points for malice, which is an element of second-degree murder, and 10 points if the intent to kill was mitigated by provocation or the conduct amounted to gross negligence. (Note: Ten points would normally be scored for manslaughter sentences, but 25 points could be awarded if the defendant pleaded down from second-degree murder to manslaughter, which carries a maximum of 15 years.)

Among murder sentences, 21.9% received 50 points and 49.8% received 25. Among assault with intent to murder sentences, 18.4% scored 50 points and 47.3% scored 25.

Aggravated Physical Abuse (OV 7). OV 7 awards 50 points if "a victim was treated with sadism, torture or excessive brutality" or conduct designed to substantially increase the victim's fear and anxiety. The Michigan Supreme Court has upheld awards of 50 points for racking a shotgun that was displayed during a carjacking and for striking store employees in the head with an airsoft gun designed to look like a sawed-off shotgun during a robbery.[25] Each person who was placed in danger of injury or loss of life is considered a victim. OV 7 was scored in 12.7% of murder sentences, 9.3% of assault sentences, 11.9% of CSC sentences, and 4.0% of robberies.

Victim Asportation or Captivity (OV 8). OV 8 awards 15 points if "a victim was asported to another place of greater danger or to a situation of greater danger or was held captive beyond the time necessary to commit the offense." As with OV 7, "each person who was placed in danger of injury or loss of life is a victim for purposes of scoring OV 8." No distinction is made for incidental or minor movements.[26] OV 8 cannot be scored when the sentencing offense is kidnapping, but since 2006, it is scored when the sentencing offense is unlawful imprisonment. Among the murder, assault, and robbery sentences, 5 to 8% were awarded 15 points. However, among CSC sentences, 15 points were awarded in nearly 18%.

Number of Victims (OV 9). OV 9 scores 100 points if more than one death occurred in a homicide case. This occurred in 3% of the murder sentences. Fewer points are scored depending on the number of victims who were placed in danger of injury or death, with 25 for 10 or more victims and 10 points for two to nine victims. No distinction is made between people who are direct subjects of the offense and responding law enforcement personnel.[27] Awards of 25 points are rare. However, the proportions of sentences that received 10 points were: murder – 28.7%, assault – 49.4%, CSC – 14.8%, and robbery – 39.8%.

Exploitation of Vulnerable Victim (OV 10). OV 10 awards 5 points if the defendant exploited an advantage in size or strength or if the victim was intoxicated, drugged, asleep, or unconscious. It awards 10 points if the defendant exploited a victim's physical or mental disability, youth or old age, a domestic relationship, or the defendant's authority status. It awards 15 points if "predatory conduct was involved." Points for OV 10 are far and away most common in CSC sentences; 61.1% received 10 points, and 23.6%

RECEIVED by MSC 4/12/2022 1:07:49 PM

received 15, even though many of the facts scored in OV 10 are elements of the offense described in 750.520(b)(1). In the other offense groups, the proportion of sentences that received zero points was nearly 79% for murder, 86% for assault, and 87% for robbery.

Criminal Sexual Penetrations (OV 11). OV 11 scores 25 points for one sexual penetration beyond the one for which the defendant is being sentenced and 50 points for two or more. It is designed to enhance punishment for multiple penetrations in the same incident *regardless of whether they are charged.* For instance, if the defendant is convicted of two counts of penetration against the same victim, 25 points can be scored in each sentencing based on the other conviction. If one of two counts is dismissed as part of a plea bargain, 25 points can be added to the one on which the defendant is sentenced. In the CSC group, 42.3% of sentences received zero points, 20.3% received 25 points, and 37.4% received 50 points. In our other three offense groups, at least 99.5% of the sentences received no points. Conduct that is scored under OV 11 cannot be scored under OV 12 and can be scored under OV 13 only in limited circumstances.

Contemporaneous Felonious Criminal Acts (OV 12). OV 12 enhances the sentence by scoring felonious conduct *not resulting in conviction* that occurred within 24 hours of the sentencing offense. It scores 25 points for three or more contemporaneous felonious acts against a person with five or 10 points for fewer or non-person crimes.

The same conduct cannot be scored in both OVs 11 and 12. However charges can be prosecuted in ways that maximize scoring under both OVs. Suppose, for instance, the defendant kidnaps the victim, strikes her, sexually penetrates her three times during a 24 hour period and resists arrest. If he pleads guilty to kidnapping, he can be given 50 points under OV 11 for two penetrations and 25 points under OV 12 for the third penetration, the assault and the resisting. The only type of criminal conduct that cannot be scored under OV 12 is the possession of a firearm during the commission of a felony (i.e., "felony firearm").

OV 12 reflects the policy of "real offense" sentencing by scoring conduct found to have occurred even though it was not used to support a conviction. It allows for the sentence to be increased based on uncharged conduct or charges dismissed in plea negotiations. If the defendant has not admitted the conduct, punishment can nonetheless be increased based on judicial fact-finding that does not require proof beyond a reasonable doubt or other procedural safeguards. *The Model Penal Code: Sentencing* would prohibit "real offense" sentencing.[28]

In each offense group, 4 to 6% of sentences received 25 points for OV 12. Roughly three-quarters of the sentences received zero points for OV 12 except in the CSC group, where the proportion of sentences with zero points was 82.0%. Presumably, multiple sexual penetrations either resulted in conviction or, if uncharged, were primarily scored under OV 11, where they would receive more points.

Continuing Pattern of Criminal Behavior (OV 13). Perhaps the most significant of the variables added by the statutory guidelines, OV 13 has multiple aims and scoring options. It is meant to further enhance the penalties for defendants with a pattern of sexual assaults on one or more children younger than 13, for defendants with substantial prior criminal conduct, and for gang-related behavior. The scoring is complex. All crimes within a five-year period, including the sentencing offense, must be counted *without regard to whether the offense resulted in a conviction* and may include juvenile adjudications. However, conduct scored for OVs 11 or 12 cannot also be scored for OV 13, unless it was gang-related.

OV 13 awards 50 points if the sentencing offense is CSC 1 and the other conduct was three or more penetrations of a child younger than 13. Twenty-five points are scored in two different situations—when the offense was part of a pattern of felonies related to inducing or maintaining gang membership, or when the offense was part of a pattern of three or more felony crimes against a person. So long as the sentencing judge finds the defendant committed multiple person crimes related to the sentencing offense, 25 points must be scored regardless of whether those crimes resulted in convictions. Other combinations of events can bring awards of 10 or five points. There is no prohibition against using the same convictions to score OV 13 and PRV 7. Thus:

RECEIVED by MSC 4/12/2022 1:07:49 PM

- PRV 7 addresses multiple counts resulting in conviction.
- OV 11 addresses multiple penetrations arising from the sentencing offense.
- OV 12 addresses multiple felonies committed within 24 hours of the sentencing offense that will not result in conviction.
- OV 13 addresses a pattern of criminal conduct committed over a period of years regardless of whether there was a conviction.

In the CSC group, 27.1% of the sentences scored 50 points; 26.6% scored 25 points; only 44.2% scored zero. Since 50 point awards are reserved for CSC sentences, virtually no sentences in the other offense groups received that many. In fact, from 63 to 80% of sentences in the other groups received zero points for OV 13. However, awards of 25 points were still substantial: murder – 13.3%, assault – 25.3%, and robbery – 27.8%.

Offender's Role (OV 14). OV 14 awards 10 points if the defendant was "a leader in a multiple offender situation." "Leader" is not defined. Judges are instructed to "consider the entire criminal transaction when determining the offender's role." The instructions specifically note that if the offense involved three or more people, more than one may be determined to have been a leader (i.e., of three co-defendants, two can be scored as leaders). In the CSC group, less than 2% of the sentences received 10 points. In the other groups, the proportion of 10-point awards ranged from 7.1% for assault to 11.5% for robbery.

Degree of Negligence Exhibited (OV 17). OV 17 is scored only if the offense involved the operation of a vehicle, vessel, ORV, snowmobile, aircraft, or locomotive. It awards five points if the defendant failed to show ordinary care and 10 points if they showed wanton or regardless disregard for life or property. It cannot be scored if points are assessed under OV 6 (intent to kill or injure). While the overwhelming majority of sentences in every offense group received no points for OV 17, 6.4% of the murder sentences and 4.4% of the assault sentences received 10.

Threat to security of Penal Institution or Court or Interference with Administration of Justice or Rendering of Emergency Services (OV 19). OV 19 scores 25 points for threatening the security of a penal institution or court, 15 points for using force or the threat of force against a person or property to interfere with the administration of justice, and 10 points for "otherwise" interfering or attempting to interfere with the administration of justice. The conduct scored usually occurs after the conviction offense was committed but not necessarily after charges were filed. For instance, 10 points may be awarded for resisting arrest and for acts meant to avoid arrest or cover up the crime, like giving the police a false name or disposing of a weapon. The conduct scored may also be unrelated to the conviction offense, such as attempting to smuggle contraband into a jail or prison. At the other extreme, the defendant can be awarded points under OV 19 even if the conduct is also the basis for the conviction offense (e.g., perjury[29] or smuggling drugs into a prison[30]) and will result in a consecutive sentence.[31]

Among all the offense groups, awards of 15 and 25 points are rare. However, awards of 10 points occurred in 3.3% of the CSC cases, 8.1% of the murder cases, 8.6% of the assault cases, and 10.2% of the robbery cases.

In sum, the offense variables were very carefully designed to enhance the sentences for second-degree murder, assault with intent to murder, and first-degree criminal sexual conduct (especially when the victim was a child) and, to a lesser degree, for armed robbery by every means imaginable. It is inherent in the structure of the guidelines that the most serious offenses fall on the grids that recommend the longest sentences precisely because of the nature of the conduct and the harm caused. However, multiple techniques are used to increase the punishment by pushing the case into cells with higher recommended minimum sentence ranges.

These techniques go well beyond awarding points for uncommon "extra" conduct, like torture, kidnapping, or causing permanent physical incapacity. Points are scored for facts that are actually required elements of the offense or that, as a practical matter, occur routinely in crimes against people.

RECEIVED by MSC 4/12/2022 1:07:49 PM

The OVs are not designed to punish just the conviction for which the defendant is being sentenced or even just the "real" offense (i.e., conduct in the same incident that supported charges which were reduced or dismissed in plea negotiations). Between OVs 11, 12, 13, and 19, along with the PRVs, the guidelines are designed to punish virtually every form of misconduct in which the defendant has engaged, whether prior to, concurrent with, or subsequent to the conviction offense, regardless of whether that misconduct resulted in a conviction or was anything like the current offense in nature or severity. The punishment is enhanced further if the conviction offense was gang-related. The same convictions can be scored as PRVs and OVs and, of course, as the basis of habitual offender enhancements. The guiding principle appears to be that once a person is convicted of a serious assaultive or sexual offense, no justification for increasing their incarceration is too much of a reach.

D. The Intersection of the PRVs and OVs on the M2 and A Grids

The recommended sentence is determined by where the PRV and OV levels intersect on the applicable grid. The manner in which the PRV and OV scores intersect demonstrates the greater impact of high-offense scores. We saw in Table 9 that, for every offense group, the greatest concentration of sentences was at PRV Levels C and D. We saw in Table 10 that, except for robberies, the greatest concentration of sentences was at the highest OV levels. Thus the OV scores appear to have more impact on pushing cases into the ranges that recommend longer minimum sentences.[32]

- For the assault group, by far the largest concentration of cases is in the C-VI (117 cases) and D-VI (174 cases) cells, where the recommended ranges begin at 135 and 171 months, respectively. Out of 36 cells on the A grid, these two account for nearly 27% of the assault sentences. Twenty cells have 20 or fewer sentences; nine of the 12 cells at OV Levels I and II have five or fewer sentences.
- For the CSC group, the largest number of cases is also in the C-VI cell. That is, despite their relatively modest prior records, these defendants also fall into higher guidelines ranges because of their offense scores. Nearly 40% of all the CSC sentences are concentrated in five cells: C-III, IV, V, and VI, and D-VI.
- For the murder group, which has only three offense levels, the largest number of cases falls in the C-III cell, with a recommended sentencing range that begins at 225 months. The second largest number of cases falls in the D-III cell, which begins at 270 months. Together, these two cells account for 27% of all the M2 sentences. Notably, there are 112 sentences in the A-III cell, with a recommended sentencing range that begins with 162 months. That is, they have zero prior record points but 100+ OV points. Conversely, there are only 90 sentences in all the OV Level I cells, regardless of PRV Level.
- By comparison, robbery cases, which tend to have lower OV scores, are concentrated in the C-II, C-III, D-II and D-III cells, where the ranges begin at 51, 81, and 108 months, respectively. Collectively these four cells account for 32% of the robbery sentences. Just 6.6% of the sentences fell into the C-V, C-VI, D-V, or D-VI cells.

For further details, see the distribution of sentences displayed in Figures 5a–5d in Chapter Eight.

RECEIVED by MSC 4/12/2022 1:07:49 PM

E. Broader Cell Ranges for Habitual Offenders

The enabling legislation for the statutory sentencing guidelines directed the Commission to:

> Establish separate sentence ranges for convictions under the habitual offender provisions in sections 10, 11, 12 and 13 of this chapter, which may include as an aggravating factor, among other relevant considerations, that the accused has engaged in a pattern of proven or admitted criminal behavior. [33]

The Commission initially chose to satisfy this charge by increasing the high end of the applicable guidelines range by 5, 10, or 15%, depending on whether the defendant was being sentenced as a second, third or fourth offender. However, at the eleventh hour, during the same meeting at which the Commission voted on the entire package of proposed guidelines, the following amendment was adopted:

> Proposal to Revisit the Treatment of Habitual Offenders and to Increase the Appropriate Cell Maxima from the Current 5%, 10%, and 15% for 2nd, 3rd, and 4th Habitual Offenders (respectively).[34]

The result of the amendment was to increase the maximum allowable minimum sentence by 25, 50, or 100% for second, third, and fourth offenders, respectively.

The decision to multiply habitual offender increases by a factor of five or more was apparently rooted in a consultant's report on a pilot project that was submitted to the Commission just six weeks earlier.[35] The project tested the impact of the proposed guidelines by having 30 judges from six counties and Recorder's Court score a sample of 344 cases.[36] Of this total, two sentences were on the M2 grid; 13 were on the A grid.[37] The sample included 31 cases of defendants being sentenced as habitual offenders—17 as habitual 2nd, five as habitual 3rd , and nine as habitual 4th. On what grids the habitual sentences fell is unclear.[38]

The sentences selected by the judges were compared to the high end of what would be the applicable guideline cell. Approximately 45% of the sentences (i.e., 14 cases) were below the cell maximum; 13% of the sentences (four cases) were at the cell maximum; 42% of the sentences (13 cases) were above the cell maximum. The extent to which the latter 13 sentences exceeded the cell maximum was:

- 1–25% – 2
- 25–50% – 1
- 51–75% – 1
- 76–100% – 3
- 101–200% – 5
- 201%+ – 1[39]

The level of habitual offender status these 13 cases involved is unclear.

The "Habitual Offender Analysis" concluded:

> As can be seen, when the judge goes above the pre-existing cell maximum, the sentence is oftentimes a 100% or more increase. Although the number of cases are quite small, it may make sense to look at the following habitual policy consideration:
>
> > Habitual 2nd – increase cell maximum by 25%
> > Habitual 3rd – increase cell maximum by 50%
> > Habitual 4th – increase cell maximum by 100%[40]

RECEIVED by MSC 4/12/2022 1:07:49 PM

In sum, the pilot project included only 31 habitual sentences, 58% of which did not exceed the high end of the guidelines cell. Why it would "make sense" to base such a consequential change in the guidelines ranges for habitual offender sentences on the remaining 13 sentences was not explained.

Using habitual offender status as an excuse to expand the sentencing guidelines range raises multiple troubling issues. Criminal history has always been an important consideration in sentencing. Long before that consideration was structured by guidelines scoring systems, individual judges gave a defendant's prior record whatever weight the judge felt it deserved in setting the *minimum* sentence. In Michigan, that was limited only by the *Tanner* rule, which says the minimum cannot exceed two-thirds of the maximum.

The habitual offender statutes allow judges to increase the *maximum* sentence beyond what is set by the penal code for the conviction offense. If the conviction offense already carries a penalty of life or any term, the habitual offender statute cannot make the potential *maximum* sentence any harsher. However, by leveraging the defendant's habitual offender status to raise the guidelines recommendation for the *minimum* sentence, defendants convicted of life-maximum offenses may also have their prison terms lengthened by the prosecutor's decision to pursue habitual offender charges.

There is no limit on the age, nature, or severity of the prior felonies that can be used to support a habitual offender enhancement. A 40-year-old defendant convicted of larceny from a person, which has a 10-year maximum, who had three non-violent offenses at age 19, such as breaking and entering a store, receiving stolen property, and unlawfully driving away a motor vehicle, could face a sentence up to life.

Under the Michigan Supreme Court's decision in *People v. Gardner*,[41] the prior convictions do not even have to arise from separate criminal incidents. Our hypothetical defendant could have helped a co-defendant break into a liquor store, accepted his share of the stolen goods, and driven off in a vehicle stolen by the co-defendant, then been convicted of all three charges when he refused to negotiate a plea. To whatever extent the maximum is raised, under the *Tanner* two-thirds rule, the minimum can be raised as well.

The defendant's criminal history is a key component of the sentencing guidelines. Identifying a recommended minimum sentence depends on the intersection of the prior record level and the offense level. The higher the prior record score, the higher the guidelines range. Why, then, should that range be expanded based on the very same consideration—the defendant's criminal history?

This use of the habitual offender enhancement does not just double-count the defendant's prior record in a general way. It has multiple specific consequences. Most apparent, it allows for such enormously long minimum sentences and such enormously wide guidelines ranges that it effectively negates any constraints on judicial discretion in sentencing people charged as habitual offenders.

Take, for instance, a defendant we will call Dennis, who was convicted of second-degree murder but whose only prior criminal activity was breaking into stores at night, a high severity felony.

- With one prior conviction and an offense severity level of III, the recommended minimum sentence is 270–450 months (22.5–37.5 years). But if Dennis is sentenced as a second offender, the range becomes 270–562 months (22.5–46.8 years).

  Thus for someone with a single prior non-violent conviction an approved minimum sentence can become nearly 50 years and the range within which the judge can select a minimum sentence may be more than 24 years wide.
- If Dennis had two prior convictions but a lower offense severity score that put him in the E-II cell, his initial sentencing range would still be 270–450 months. But if he is sentenced as a third offender the range becomes 270–675 months (22.5–56.25 years).
- If Dennis had three prior convictions for breaking and entering but an even lower offense severity score that put him in the F-I cell, his initial sentencing range would again be 270–450 months. But if he is sentenced as a fourth offender, the range becomes 270–900 months (22.5–75 years).

RECEIVED by MSC 4/12/2022 1:07:49 PM

Equally important:

- The habitual enhancement and the PRV score can count exactly the same prior crimes. That is, the same priors can put the defendant into a higher range in the first place, and then be used to raise the high end of that range.
- Conversely, the habitual enhancement can be used to increase a minimum sentence based on priors that the sentencing guidelines weigh less heavily or exclude altogether. While the guidelines take a more refined approach, allotting points based on the severity of the prior offense and awarding no points for stale priors under the "10-year gap rule," increasing the recommended minimum sentence based on the habitual enhancement circumvents that approach by allowing any old prior conviction to push up the minimum sentence.
- Since filing habitual offender enhancements lies solely in the prosecutor's discretion, widely divergent practices across counties adds to the disparities in sentencing the guidelines were meant to reduce.

F. Bringing It All Together: Regression Analyses

The regression analyses found in Appendix B allow us to examine how the independent variables of offense type, PRV level, OV level, and habitual status contribute to sentence length.[42] In the full model that combines the four offense types, which offense was committed explains more than 25% of the differences observed in sentence length. Offense is followed by the two other variables we would expect to have substantial explanatory power, the offense severity (OV level) and prior criminal history (PRV level). These elements explain an additional 19.6%, and 23.1%, respectively. Habitual status contributes only 2.6% of the variance explained, largely because the prior convictions required for habitual status are already counted in the prior record level.[43] Together, these four variables explain more than 70% of the variance in sentence length, with about 30% left unexplained by guidelines factors. These results are similar to those of NCSC, which found that the guidelines elements account for 67% of the variation in sentence length, and that conviction offense had the greatest impact. The extent to which variance is explained by departures from the guidelines recommendations will be explored in Chapter Seven.

We can also use the regression results to estimate how the length of the sentence would change as individual guidelines factors change.

- In the full model, the increase from PRV Level A to Level C was associated with a 73% increase in sentence length, and an increase from A to F was found to have a 228% increase.
- Similarly, an increase from OV Level I to Level III was associated with a 78% increase in sentence length, and an increase from I to VI was associated with an increase of 226%.
- When all the variables were held constant as they are in the regression analyses, the increase in sentence length from a non-habitual to a second, third, and fourth habitual was 10%, 19%, and 25%, respectively.

Breaking the full model into separate analyses based on the offense committed allows us to compare the relative importance of the guidelines factors to see if there are differences among offenses. As we have found previously, examining the life-maximum sentences together can mask important variations. Since each of these regressions is focused on a specific offense, the first variable included is the OV level, which is found to have differing amounts of explanation across the offenses.

- Offense severity explains a minimum of only 5.4% of the variance in sentence length for second-degree murder offenses and a maximum of 29.3% in armed robberies.
- Prior record explains 19% for second-degree murder to 36% for armed robberies of the variance in sentence length.

RECEIVED by MSC 4/12/2022 1:07:49 PM

- These results suggest at least two things. First, murder is so inherently different than other crimes that the nature of the offense overwhelms the importance of other guidelines factors. Second, because murder is different it is on its own grid with higher recommended sentences that itself explains much of the variance in sentence length.
- Like the full model habitual status does not explain a large amount of the difference in sentence lengths, only 5% or less.

When we examine the amount of variance the guidelines factors explain for each offense type, the results are quite different than they were for the full model. For murder sentences, the four sentencing guidelines factors explain only 29%. For assault with intent, it is 53%, for CSC it is 60%; and for armed robbery, they explain almost 68% of the variance in sentence length.

The estimation models separated by offense type also show different patterns for the impact of OV and PRV level and habitual status. For a second-degree murder sentence, the increase from a PRV Level A to a PRV Level C was associated with a 32% increase in sentence length—less than half of the impact observed in the full model. In contrast, the same change in PRV level brought a 91% increase for armed robbery sentences. Overall, PRV level had the largest impact on armed robbery sentences, then CSC, assault, and the least on murder. There are also differences in the impact of OV level on sentence length depending on the sentencing offense.

OV level has the most impact on sentences for CSC, with increases from OV Level I to III resulting in a 163% increase, and from Level I to VI in a 277% increase. Sentences for murder are less altered by changes in OV Level, with the increase from Level I to II resulting in an increase of 22%, and to level III in 49%. The impact of habitual status is also the greatest for CSC offenses, ranging from an increase of 17% for a second habitual offender, 29% for third offender, and 34% for fourth.

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER SEVEN. HOW THE GUIDELINES PROMOTE DISPARITY

The offense determines the sentencing grid on which the case belongs. The intersection of the points awarded for the prior record and offense variables determines which cell on that grid contains the range of months within which the defendant's minimum sentence should be chosen. Thus, the PRV and OV scores set the parameters for the recommended sentence. However, these parameters provide only a starting point for determining the actual sentence.

A. How Sentences are Distributed Within, Above and Below the Applicable Range

Recall that the legislative mandate was to "Reduce sentencing disparities based on factors other than offense characteristics and offender characteristics and ensure that offenders with similar offense and offender characteristics receive substantially similar sentences." Yet the legislative guidelines, like their judicial predecessor, were deliberately designed to permit similarly situated people to receive very different punishments. These disparities arise in two ways.

First, the ranges are very broad. Judges are free to select a sentence within these parameters for any reasons they choose. The Sentencing Commission said it sought "to create a balance between unfettered judicial discretion and confining discretion too stringently and thereby destroying the goal of individualized sentences."[1] The M2 and A-D grids are designed so the low and high ends of the cells represent a variance of 25%, plus or minus, from the cell's midpoint.[2] Because the sentences at the midpoint of the cells on the M2 and A grids are so long, this variance creates extremely wide ranges.

Take, for instance, someone whose offense is on the A grid, whose prior record Level is D, whose offense variable Level is V, and who is not being sentenced as a habitual offender. The cell range is 135–225 months. That is 90 months or 7.5 years. While the midpoint of the range is 180 months, or 15 years, two people with identical PRV and OV scores could get sentences as different as 11.25 and 18.75 years and still be within the guidelines recommendation.

On the M2 grid, even people who have the lowest possible scores on the both the offense and prior record variables could, without a departure, receive minimum sentences that differ by five years. At the other extreme, in the F-III cell of the M2 grid, the distance between the low and high ends of the recommended range is 20 years. As Sentencing Commission Chair Judge Paul Maloney observed in reference to the judicial guidelines: "Frankly, I find myself at some point wondering whether the guideline cells, with that breadth of ranges, are providing much guidance at all as to what sentence I ought to impose."[3]

Second, even with such broad choices available, judges have always been permitted to depart from the ranges by setting sentences that are below the low end (the minimum-minimum) or above the high end (the maximum-minimum). In 1994, the overall compliance rate with the advisory second judicial guidelines was 88%.[4] Compliance with the legislative guidelines was mandatory until 2015, but departures were permitted so long as an appellate court agreed they were supported by "substantial and compelling reasons."

The Sentencing Commission, which differentiated only by grid, not by offense, anticipated that, for the M2 grid, the rate of upward departures would be 12% and downward departures would be 15%, leaving a compliance rate of 73%. For the A grid, the Commission anticipated a compliance rate of 80%, with 5% upward departures and 15% downward departures.[5]

This projection proved to be far too optimistic. Compliance rates vary greatly by guidelines grid. The large majority of sentences fall onto grids where the ranges are narrower because the statutory maximum sentences are lower. When the maximum sentence is five or 10 years, and the minimum cannot exceed two-thirds of the maximum, there is simply a lot less room to move. The National Center for State Courts (NCSC) found that, in 2004, compliance rates for grids D–H were between 94 and 99%; for the C grid, compliance was nearly 82%.

However, for the M2, A, and B grids, NCSC found compliance rates between 55 and 58%, with downward departures being far more frequent than upward departures. The authors concluded: "Given the

RECEIVED by MSC 4/12/2022 1:07:49 PM

wide bandwidth, this level of downward departure sends a message that Michigan judges believe that the recommended sentences are too high."[6]

Table 11 shows that our findings on the life-maximum sentences were similar to those of NCSC, with compliance rates for non-habitual sentences ranging from 51.0 to 63.2%. Table 11 also shows how departures in non-habitual sentences under the legislative guidelines compared to departures under the predecessor judicial guidelines. Downward departures increased for every offense, more than doubling for murder sentences and nearly quadrupling for robbery sentences. Upward departures also increased, albeit much more modestly. The combined result was a decline in compliance rates that ranged from 10 points for CSC sentences to 31 points for armed robbery sentences.

Table 11. Percentage of Departures under 2nd Judicial and Legislative Guidelines (Non-Habitual sentences)[7]

Crime	Downward Departure		Within Range		Upward Departure	
	Judicial 2nd	Statutory	Judicial 2nd	Statutory	Judicial 2nd	Statutory
Murder, 2nd	12.9	32.3	78.1	56.0	9.0	11.7
Assault	23.9	38.6	70.7	51.0	5.4	10.7
CSC, 1st	19.2	23.3	73.2	63.2	7.5	13.6
Robbery	10.5	40.3	85.4	54.3	4.4	5.4

Ironically, the mandatory legislative guidelines actually resulted in much less compliance than the advisory judicial guidelines, which were designed to reflect actual sentencing practices. The Sentencing Commission chose a policy of dramatically increasing sentences for life-maximum offenses. The judges who were expected to enforce that policy often chose not to follow it.

1. Non-Habitual Sentences

Table 12 shows us that, overall, 4,008 of the 8,986 non-habitual sentences, or 44.6%, were either upward or downward departures; thus, only 55.4% followed the guidelines recommendations. Table 12 also allows us to see how many sentences were right at the lower and upper margins of the guidelines ranges, that is, at either the "minimum-minimum" or the "maximum-minimum." Because they were required to state "substantial and compelling reasons" for departing from the recommended ranges, some judges may have chosen to go only as far as selecting a sentence at the limits of the ranges within which they still had complete discretion. Given the breadth of the guidelines ranges, adding the cases at the margins to the departures gives a strong sense of how many sentences within each offense group were relatively long and relatively short.

- For murder, 38.3% of the sentences were at the low end and 19.7% were at the high end.
- For assault, 47.6% were at the low end and 16.7% were at the high end.
- For CSC, 33.4% were at the low end and 25.0% were at the high end.
- For robbery, 59.0% were at the low end and 8.5% were at the high end.

All told, for the life-maximum offenses, only about a third are sentenced somewhere along the continuum of sentences within the recommended ranges, as opposed to at, below, or above the margins.

A closer look at the sentences between the extremes shows how often judges go straight to the middle of the applicable range as opposed to making a deliberate decision to go toward the low or high end. Overall, just 6.3% of the non-habitual sentences fell directly at the midpoints of their respective cell ranges. From half to three-quarters of the non-habitual sentences in every offense group fell somewhere below the midpoint:

RECEIVED by MSC 4/12/2022 1:07:49 PM

- Murder – 56.9%
- Assault – 63.6%
- CSC – 52.7%
- Robbery – 77.9%[8]

Because the large volume of robberies so heavily impacts the total, overall more than two-thirds of the life-maximum sentences were below the midpoint of the applicable guidelines cells and more than one-third were below the minimum-minimum. This finding confirms the NCSC authors' conclusion that Michigan judges see the guidelines recommendations as being too high.

Sentencing judges clearly treat the offenses on the A grid differently from each other. Armed robbery had the greatest proportion of downward departures and the smallest number of upward departures. Conversely, sex offenses had a far lower rate of downward departures and more upward departures. To the extent that judges think the guidelines recommendations are too high, they feel less so with regard to people sentenced for sex offenses.

Table 12. Percentages of Minimum Sentences across Guidelines Ranges, by Habitual Level, 2003–2012

	N	Median (in mos.)	Below Min-Min	At Min-Min	Below Midpoint	At Midpoint	Above Midpoint	At Max-Min	Above Max-Min
M2									
None	1,218	240	32.3	6.0	18.6	11.4	12.2	8.0	11.7
2nd	73	360	17.8	5.5	12.3	9.6	13.7	6.8	34.2
3rd	50	368.5	19.6	9.8	21.6	2.0	11.8	2.0	33.3
4th	58	540	15.8	5.3	14.0	1.8	8.8	5.3	49.1
AWIM									
None	884	132	38.6	9.0	16.0	5.4	14.3	6.0	10.7
2nd	67	240	13.4	14.9	13.4	1.5	9.0	0.0	47.8
3rd	53	300	1.9	5.7	17.0	15.1	5.7	1.9	52.8
4th	85	300	12.9	9.4	10.6	7.1	7.1	1.2	51.8
CSC									
None	1,225	126	23.3	10.1	19.3	8.1	14.3	11.4	13.6
2nd	118	232.5	6.0	8.5	9.4	4.3	18.8	5.1	47.9
3rd	64	336	0.0	4.6	6.2	4.6	20.0	1.5	63.1
4th	103	300	5.8	1.7	12.6	4.9	11.7	1.0	53.4
RA									
None	5,659	72	40.3	18.7	18.9	5.0	8.6	3.1	5.4
2nd	408	126	15.1	22.2	21.2	5.6	13.9	2.9	19.0
3rd	288	165	16.5	12.7	18.7	4.6	14.4	2.5	30.6
4th	531	180	12.0	17.1	15.6	6.4	11.8	2.3	34.9
TOTALS									
None	8,986	96	36.7	14.9	18.6	6.3	10.4	5.2	7.9
2nd	666	169.5	13.7	17.3	17.3	5.4	14.4	3.5	28.5
3rd	455	216	13.0	10.3	17.1	5.3	13.6	2.2	38.5
4th	777	224	11.5	14.5	14.5	6.0	11.1	2.2	40.2
OMNI Data, All offense types significant at $p < .001$									

Table 12 shows how sentences were distributed across the range of choices that judges can make once the applicable cell has been identified. But what do these choices mean as a practical matter? How long are the sentences that are above and below the extremes? How long are those within the guidelines range but not right at either end?

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 13 answers these questions. It looks at the single most commonly used cell for each offense type, then displays the median length of the sentences that fell at each point on the range. So, for example, the cell in which murder sentences most commonly fall is C-III, which has a min-min of 225 months and a max-min of 375 months. For the 58 sentences imposed with downward departures, the median sentence was 180 months or 15 years. For the 14 sentences imposed with upward departures, the median sentence was 480 months or 40 years. Thus, for the defendant the choice between an upward and a downward departure meant 25 years of incarceration. Indeed the difference between the median sentences at and above the max-min was nearly nine years.

For assault sentences in the D-VI cell, the difference between the median downward departure and the median upward departure was 19.75 years. For CSC sentences in the C-VI cell, the difference was 17 years. Because the most common cell for robbery sentences was C-II (i.e., at much a lower OV level than the other offenses), the difference between the median downward departure and the median upward departure was six years.

Table 13. Median Sentence by Place on the Range for Most Common Cells per Offense Non-Habitual, 2003–2012 (Number of sentences in parentheses)

	Below Min-Min	At Min-Min	Below Mid-point	At Mid-point	Above Mid-point	At Max-Min	Above Max-Min
Murder/Cell C-III							
Median (197)	180 (58)	225 (6)	240 (50)	300 (25)	360 (35)	375 (9)	480 (14)
Assault/Cell D-VI							
Median (139)	120 (46)	171 (5)	196 (28)	228 (4)	240 (29)	285 (13)	357 (14)
CSC/Cell C-VI							
Median (169)	96 (26)	135 (9)	150 (37)	180 (31)	216 (23)	225 (18)	300 (25)
Robbery/Cell C-II							
Median (669)	36 (231)	51 (167)	60 (134)	68 (8)	78 (85)	85 (9)	108 (35)
OMNI Data, All offense types significant at $p < .001$							

The relationship of place on the range to disparity in sentencing is amplified when we examine redundant cells (i.e., different cells with the same recommended ranges). Consider the four cells on the A grid [C-VI, D-V, E-IV, F-III] with a range of 135–225 months (11.25–18.75 years) and midpoint of 180 months (15 years). From 2003–2012, 223 people convicted of sex offenses who were not charged as habitual offenders fell into these cells. Of these:

- 43 received minimums below 11.25 years, the lowest being 3.5 years
- 12 were sentenced to exactly 11.25 years,
- 46 received sentences above 11.25 and below 15 years,
- 39 were sentenced right at the midpoint to 15 years,
- 28 received sentences greater than 15 but less than 18.75 years,
- 21 were sentenced to exactly 18.75 years,
- 34 received minimums above 18.75 years, the highest being 65 years

RECEIVED by MSC 4/12/2022 1:07:49 PM

Even assuming the cases at the greatest extremes were unique, this distribution, which is more or less replicated in every cell for every offense, shows that large disparities in the sentences imposed on people with similar prior records and similar offense severity are routine.

2. Habitual Sentences

Returning to Table 12, we also find the minimum terms imposed on people sentenced as habitual offenders. This allows us to compare the distribution of habitual sentences across guidelines ranges with non-habitual sentences and among habitual sentences depending on the defendant's status as a second, third, or fourth offender. In particular, we can see to what extent judges actually use their authority to exceed the high end of the guidelines ranges based on the defendant's habitual offender status. The major takeaways are the following:

- Not surprisingly, at 12.6%, the rate of downward departures was far lower for habitual than for non-habitual sentences. Depending on the offense group and the habitual level, the rate of downward departures was two-three times lower for habituals. The differences were even more extreme for the CSC group.
- What *is* surprising is how many downward departures there were among habitual sentences. In the murder group, more than 15% of the habitual sentences were below the low end of the guidelines range. In the robbery group, it was more than 12%. Except for those sentenced as third habitual offenders it was roughly 13% for assault sentences. Only for the CSC group were downward departures for habitual sentences at 6% or below.

  Moreover, when we look at how many habitual sentences were below the midpoint, we find:

  Habitual second – 48.3%
  Habitual third – 40.4%
  Habitual fourth – 40.5%

That is, despite the designation of habitual offender status—even fourth habitual status—in a substantial proportion of cases judges chose to impose sentences at the lower end of the applicable guidelines range. Indeed, that designation did not prevent downward departures in more than 10% of habitual sentences at every level.

- Again, not surprisingly, the rate of sentences above the high end of the standard guidelines range was from three and a half to six times higher for habitual than for non-habitual sentences, depending on the offense group and habitual level. Nonetheless, the overall percentage of sentences above the maximum-minimum was just:

  Habitual second – 28.5%
  Habitual third – 38.5%
  Habitual fourth – 40.2%

- Here, too, the surprise is that things are not more extreme. Even with the large difference from non-habitual sentences, with just a few exceptions, only about one-third to one-half of the habitual sentences for every offense at every habitual level exceeded the standard maximum-minimum. Specifically the percentage of each offense group was:

RECEIVED by MSC 4/12/2022 1:07:49 PM

Murder – 38.7%
Assault – 50.7%
CSC – 53.5%
Robbery – 28.2%

- The differences between habitual and non-habitual sentences at different points along the guidelines range are more varied and much less extreme than they are at the far ends.

It is quite clear that regardless of offense or habitual status, judges do not feel compelled to exercise their authority to impose on habitual offenders minimum sentences that are 25, 50, or 100% higher than the maximum-minimum applicable to defendants who were not habitualized. In fact, of 1,898 enhanced sentences that were at any habitual level, only 35.7% (677 sentences) exceeded the basic max-min. Moreover, when we examine how far above the max-min these 677 sentences were, we find the increases were not typically as high as they could have been.

We took all of the enhanced sentences that exceeded the max-min and determined the percentage by which each sentence was greater. For each offense group and each habitual level, we rank ordered the extent of the excess, which allowed us to find the median percentage above the basic max-min. Looking at each offense group regardless of habitual offender level, the percentages above the max-mins were:

Murder – 28.0%
Assault – 26.3%
CSC – 31.4%
Robbery – 26.3%

That is, despite the allowable increases of 25, 50, and 100%, the median extent of the increases hovered close to 28%.

When we look at each level of habitual status regardless of offense group, the median extent above the basic max-mins was:

Habitual second – 23.3%
Habitual third – 33.3%
Habitual fourth – 33.3%

Thus, from the perspective of both offense type and habitual level, the median extent of the increases was closer to 25% than to the 50% allowed for habitual third or the 100% allowed for habitual fourth.

Of course, by definition, medians tell us only what is in the middle, not what is at the extremes. Habitual offender sentences that did exceed the basic max-min ran the gamut from slightly above to those that departed well above even the 100% increase in the recommended range that the guidelines allow for fourth offenders. Certainly, the availability of increased ranges for defendants sentenced as habitual offenders makes a great deal of difference in individual cases and pushes sentences over the high end of the normal range more than a third of the time. However, it is apparent that judges exercise a great deal of individualized discretion. Perhaps they feel less need to utilize their expanded authority because the base ranges are already broad enough to provide them with abundant discretion, and the guidelines scoring provides ample consideration of the defendant's prior record.

Finally, we turn back for a moment to the differences among sentences at different habitual levels. We saw in Chapter Five, Figure 4, that the biggest impact on the median sentence occurs when the sentence goes from being non-habitual to habitual second. The changes thereafter, from habitual second to habitual third and from habitual third to habitual fourth, are generally much less substantial. Table 12 also indicates that the differences between each habitual level are often not great and that there are generally not clear progressions.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Take, murder, for example. The differences in the proportion of habitual sentences that are below the min-min is small, and the habitual third group has the largest share. At the other extreme, the habitual second and third groups have nearly identical proportions of sentences above the basic max-min with a big increase between habitual third and fourth. By comparison, the CSC group has no habitual third sentences below the min-min but the largest share by far of sentences above the max-min. Overall, there is little difference among the habitual levels in terms of proportion of sentences below the min-min but a substantial difference between habitual second and habitual third among sentences above the max-min.

3. Departures in the Regression Analyses

In the full model, the addition of departures below and above the guidelines range explains an additional 18% of the variance in sentence length, bringing the cumulative total explained to 88.8%. Considering that the sentencing guidelines factors have already established very broad recommended ranges, having departures—which are essentially a rejection of those ranges—explain almost one-fifth of the variance is concerning.

While the inclusion of departures was an important factor in the overall model, the inclusion of departures when examining the offenses separately displays important differences. Fully half of the variance in M2 sentences can be explained by departures, compared to 30% of assault, 26% of CSC sentences, and less than 20% of the variance in armed robbery sentences.

Despite the range of explanatory power for variation in sentence length, the actual impact of departures on sentence length is fairly consistent. In the full model, if a sentence were a downward departure instead of being within the guidelines range, we estimate a 44% decrease in sentence length, and a 76% increase if the sentence were an upward departure. Notably, while NSCS found a greater estimated increase in sentence length for upward departures, their downward departure estimated decrease of 48.5% was practically identical.

In the offense-specific models, we see less variation in impact than we observed with other variables. The impact of a downward departure ranged from a 46% decrease for CSC to a 43% decrease for robbery. The impact of upward departures is a bit more varied, ranging from a 59% increase for murder sentences to a 82% increase for CSC sentences.

B. The Impact of PRVs and OVs on Where Within Ranges Sentences Fall

So what, apart from habitual status, determines where on the broad continuum of possibilities any particular sentence will fall? Several possible factors may influence how judges exercise the very broad discretion the guidelines afford them in life-maximum cases.

- Judges may give more or less weight to the individual and offense characteristics identified by the PRVs and OVs.
- Sentences may be affected by systemic factors such as the county of conviction or whether the defendant pleaded guilty as opposed to having a jury trial.
- Sentences may be influenced by identifying characteristics of the defendant, such as age, race, or sex, or by individualized information about the defendant's employment status, family history, mental health, or substance abuse.
- Individual judges may have personal predilections about what sentences are appropriate for particular crimes.

We will first explore the possibility that PRV and OV scores systematically explain how judges choose specific sentences within or outside the recommended range. We will examine the influence of extra-guidelines factors in subsequent chapters.

RECEIVED by MSC 4/12/2022 1:07:49 PM

1. PRV Levels

If judges' decisions about sentences are heavily influenced by the defendants' criminal histories, we would expect to see lower PRV levels yield more downward departures and higher PRV levels yield more upward departures; however, the data for the non-habitual sentences show that this expectation is not accurate.

- For the murder group, 30% of those who received upward departures scored at the lowest PRV Level, A. Only 4% scored at the highest level, F. Conversely, a quarter of those who received downward departures were at the two lowest PRV levels, A and B; another quarter were at the two highest levels, E and F.
- Upward departures among assault sentences have 34% scoring at PRV Level A and just 9.5% scoring at the highest level, F. Downward departures had 25.5% scoring at the two highest PRV levels, E and F.
- Among CSC defendants, a fifth of those who received upward departures scored at PRV Level A and fewer than 5% scored at Level F. Nearly one-fifth of those who received downward departures were at Level A while another one-fifth were at the two highest levels, E and F, combined.
- For the armed robbery group, nearly 40% of those with upward departures scored at PRV Level A while only 5% scored at Level F. Among those who received downward departures, 10% scored at Level A and 10% scored at Level F.

In sum, in every offense group, people with high prior record scores received downward departures, and people with low prior record scores received upward departures. Thus, it appears that PRV level does not explain why judges choose to depart from the guidelines recommendation in either direction for any offense type. Rather, the extent to which sentences at each PRV level fall below, within, and above the recommended guidelines range tends simply to track, albeit imperfectly, the frequency with which sentences at each PRV level appear in each offense type overall.

2. OV Levels

If judges' decisions about sentences are heavily influenced by the details of the offense, we would expect to see lower OV levels yield more downward departures and higher OV levels yield more upward departures. However, the data indicate that this expectation is also not accurate. Except for the robbery group, the single largest share of both downward and upward departures are sentences that score at the highest OV levels.

- For the murder group, 57% of the downward departures compared to 44% of the upward departures were sentences scored at OV Level III, the highest on the M2 grid. Conversely, 18% of the upward departures were at OV Level I.
- Among assault sentences, 40% of the downward departures and 45% of the upward departures had OV scores at Level VI, the highest level on the A grid; however, 15% of the upward departures scored at OV Level I.
- In the CSC group, 34% of the downward departures and 30% of the upward departures scored at OV Level VI; however nearly a quarter of the upward departures were sentences that scored at the two lowest OV levels, I and II, combined.
- The robbery group shows the opposite pattern. Over 60% of the upward departures are sentences with low OV scores as are 43% of the downward departures. High OV levels account for only 12% of the downward departures and 9% of the upward departures. As we

saw in Table 10, robbery sentences are far more likely overall to have low OV scores than is the case in the other offense groups.

As with PRV levels, the frequency with which sentences at each OV level appear at each point in the sentencing range tends to reflect the frequency with which sentences at each OV level appear in each offense group overall. Having a low OV score does not increase the likelihood of a downward departure nor does having a high OV score increase the likelihood of an upward departure. On the contrary, except for the robbery group, well over half the downward departures occurred in sentences with high OV levels. In fact, for every offense type, upward departures have a much greater proportion of sentences with low OV scores than do downward departures.

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER EIGHT. HOW THE ACTUAL MINIMUM SENTENCES RELATE TO THE RECOMMENDED RANGES

So where does this complex process of grid selection, prior record, and offense variable scoring, judicial choices made within guidelines ranges, and upward and downward departures lead? In the 8,986 non-habitual sentences imposed for our four life-maximum offenses between 2003 and 2012, it led to the distribution of median sentences displayed in Figures 5a–5d. Each figure is a guidelines grid that has the median sentence for every cell on the grid written above the recommended minimum-minimum. The number of sentences that fell into that cell is written in parentheses below the recommended min-min. We can see the impact of the guidelines on proportionality and disparity by examining a half-dozen characteristics common to all the grids.

Figure 5a. Sentencing Grid for Class M2 —MCL 777.61
Includes Ranges Calculated for Habitual Offenders (MCL 777.21(3)(a)-(c)

Second-Degree Murder

OV Level	PRV Level A 0 Points		B 1-9 Points		C 10-24 Points		D 25-49 Points		E 50-74 Points		F 75+ Points		Offender Status
I 0-49 Points	168	150	216	240	174	270	240	300/L	252	375/L	215	450/L	
	90	187	144	300	162	337	180	375/L	225	468/L	270	562/L	HO2
		225		360		405		450/L		562/L		675/L	HO3
	(14)	300	(13)	480	(24)	540	(21)	600/L	(8)	750/L	(2)	900/L	HO4†
II 50-99 Points	180	240	216	270	240	300/L	240	375/L	240	450/L	315	525/L	
	144	300	162	337	180	375/L	225	468/L	270	562/L	315	656/L	HO2
		360		405		450/L		562/L		675/L		787/L	HO3
	(115)	480	(69)	540	(133)	600/L	(102)	750/L	(39)	900/L	(27)	1050/L	HO4†
III 100+ Points	216	270/L	216	300/L	264	375/L	300	450/L	330	525/L	306	600/L	
	162	337/L	180	375/L	225	468/L	270	562/L	315	656/L	365	750/L	HO2
		405/L		450/L		562/L		675/L		787/L		900/L	HO3
	(112)	540/L	(81)	600/L	(197)	750/L	(147)	900/L	(64)	1050/L	(50)	1200/L	HO4†

RECEIVED by MSC 4/12/2022 1:07:49 PM

Figure 5b. Sentencing Grid for Class A Offenses—MCL 777.62
Includes Ranges Calculated for Habitual Offenders (MCL 777.21(3)(a)-(c)
Assault with Intent to Murder

OV Level	PRV Level A 0 Points		B 1-9 Points		C 10-24 Points		D 25-49 Points		E 50-74 Points		F 75+ Points		Offender Status
I 0-19 Points	180	35	29	45	60	70	–	85	150	135	–	180	
	21	43	27	56	42	87	51	106	81	168	108	225	HO2
		52		67		105		127		202		270	HO3
	(14)	70	(1)	90	(1)	140	–	170	(1)	270	–	360	HO4†
II 20-39 Points	27	45	22	70	53	85	96	135	63	180	156	210	
	27	56	42	87	51	106	81	168	108	225	126	262	HO2
		67		105		127		202		270		315	HO3
	(5)	90	(1)	140	(6)	170	(7)	270	(2)	360	(5)	420	HO4†
III 40-59 Points	48	70	66	85	69	135	108	180	126	210	200	225	
	42	87	51	106	81	168	108	225	126	262	135	281	HO2
		105		127		202		270		315		337	HO3
	(19)	140	(13)	170	(18)	270	(21)	360	(7)	420	(5)	450	HO4†
IV 60-79 Points	67	85	81	135	96	180	120	210	120	225	120	285	
	51	106	81	168	108	225	126	262	135	281	171	356	HO2
		127		202		270		315		337		427	HO3
	(34)	170	(13)	270	(59)	360	(40)	420	(29)	450	(29)	570	HO4†
V 80-99 Points	84	135	108	180	126	210	135	225	198	285	160	375/L	
	81	168	108	225	126	262	135	281	171	356	225	468/L	HO2
		202		270		315		337		427		562/L	HO3
	(31)	270	(25)	360	(46)	420	(45)	450	(16)	570	(17)	750/L	HO4†
VI 100+ Points	120	180	126	210	180	225	210	285	286	375/L	267	450/L	
	108	225	126	262	135	281	171	356	225	468/L	270	562/L	HO2
		270		315		337		427		562/L		675/L	HO3
	(29)	360	(19)	420	(113)	450	(139)	570	(50)	750/L	(44)	900/L	HO4†

RECEIVED by MSC 4/12/2022 1:07:49 PM

Figure 5c. Sentencing Grid for Class A Offenses—MCL 777.62

Includes Ranges Calculated for Habitual Offenders (MCL 777.21(3)(a)-(c)

Criminal Sexual Conduct, First-Degree

OV Level	PRV Level												Offender Status
	A 0 Points		B 1-9 Points		C 10-24 Points		D 25-49 Points		E 50-74 Points		F 75+ Points		
I 0-19 Points	60 21 (13)	35 43 52 70	31 27 (4)	45 56 67 90	45 42 (8)	70 87 105 140	61.5 51 (2)	85 106 127 170	60 81 (2)	135 168 202 270	168 108 (2)	180 225 270 360	 HO2 HO3 HO4†
II 20-39 Points	36 27 (36)	45 56 67 90	45 42 (23)	70 87 105 140	61 51 (58)	85 106 127 170	96 81 (21)	135 168 202 270	150 108 (11)	180 225 270 360	156 126 (7)	210 262 315 420	 HO2 HO3 HO4†
III 40-59 Points	49.5 42 (46)	70 87 105 140	60 51 (15)	85 106 127 170	108 81 (92)	135 168 202 270	120 108 (25)	180 225 270 360	168 126 (21)	210 262 315 420	138 135 (6)	225 281 337 450	 HO2 HO3 HO4†
IV 60-79 Points	60 51 (60)	85 106 127 170	90 81 (23)	135 168 202 270	132 108 (114)	180 225 270 360	174 126 (44)	210 262 315 420	180 135 (20)	225 281 337 450	196 171 (12)	285 356 427 570	 HO2 HO3 HO4†
V 80-99 Points	96 81 (41)	135 168 202 270	139.5 108 (14)	180 225 270 360	156 126 (93)	210 262 315 420	166.5 135 (28)	225 281 337 450	169.5 171 (14)	285 356 427 570	168 225 (12)	375/L 468/L 562/L 750/L	 HO2 HO3 HO4†
VI 100+ Points	120 108 (49)	180 225 270 360	180 126 (10)	210 262 315 420	180 135 (169)	225 281 337 450	186 171 (78)	285 356 427 570	228 225 (28)	375/L 468/L 562/L 750/L	294 270 (24)	450/L 562/L 675/L 900/L	 HO2 HO3 HO4†

RECEIVED by MSC 4/12/2022 1:07:49 PM

Figure 5d. Sentencing Grid for Class A Offenses—MCL 777.62
Includes Ranges Calculated for Habitual Offenders (MCL 777.21(3)(a)-(c)
Armed Robbery

OV Level	PRV Level A: 0 Points		B: 1-9 Points		C: 10-24 Points		D: 25-49 Points		E: 50-74 Points		F: 75+ Points		Offender Status
I 0-19 Points	24	35	27.5	45	42	70	51	85	81	135	108	180	
	21	43	27	56	42	87	51	106	81	168	108	225	HO2
		52		67		105		127		202		270	HO3
	(151)	70	(92)	90	(165)	140	(138)	170	(72)	270	(55)	360	HO4†
II 20-39 Points	27	45	42	70	51	85	79.5	135	108	180	120	210	
	27	56	42	87	51	106	81	168	108	225	126	262	HO2
		67		105		127		202		270		315	HO3
	(417)	90	(270)	140	(669)	170	(396)	270	(160)	360	(138)	420	HO4†
III 40-59 Points	42	70	51	85	81	135	108	180	126	210	135	225	
	42	87	51	106	81	168	108	225	126	262	135	281	HO2
		105		127		202		270		315		337	HO3
	(188)	140	(157)	170	(500)	270	(419)	360	(178)	420	(139)	450	HO4†
IV 60-79 Points	51	85	68	135	108	180	120	210	135	225	156	285	
	51	106	81	168	108	225	126	262	135	281	171	356	HO2
		127		202		270		315		337		427	HO3
	(66)	170	(35)	270	(238)	360	(205)	420	(108)	450	(74)	570	HO4†
V 80-99 Points	72	135	81	180	126	210	135	225	174	285	169.5	375/L	
	81	168	108	225	126	262	135	281	171	356	225	468/L	HO2
		202		270		315		337		427		562/L	HO3
	(19)	270	(19)	360	(74)	420	(99)	450	(39)	570	(32)	750/L	HO4†
VI 100+ Points	120	180	102	210	156	225	180	285	180	375/L	240	450/L	
	108	225	126	262	135	281	171	356	225	468/L	270	562/L	HO2
		270		315		337		427		562/L		675/L	HO3
	(25)	360	(19)	420	(129)	450	(100)	570	(57)	750/L	(23)	900/L	HO4†

RECEIVED by MSC 4/12/2022 1:07:49 PM

A. The Number of Different Medians

For each offense group, the same median sentences are repeated at different points on the grid. That is, different combinations of offense level and prior record level sometimes produce the same median and sometimes produce different ones. As a result, for every offense group, the total number of different median sentences is less than the total number of cells on the grid.

- Murder – 18 cells, 12 medians
- Assault – 36 cells, 26 medians
- CSC – 36 cells, 25 medians
- Robbery – 36 cells, 19 medians

We will see that some, but not all, of the repeated medians are due to the number of redundant cells on the grids. At the least, this arrangement raises questions about what is gained by having so many cells that produce similar results.

B. Medians Below the Minimum-Minimum

Particularly noteworthy is that every offense group has medians that are actually below the min-min. Bear in mind that the median represents the midpoint of the sentences in the relevant cell and that multiples of the same sentence may cluster around the median. This means that at least half the sentences in that cell were equal to or below the median and thus departed below the low end of the recommended range. The figures show how many cells with medians below the min-min each offense group had:

- Murder – 3
- Assault – 9
- CSC – 3
- Robbery – 11

We have already seen that a substantial proportion of sentences in each offense group were downward departures; that is, they were below the min-min. This could occur, of course, even if the median of all the sentences in a particular cell was above the min-min. (For instance, if there are 100 sentences in a cell with a min-min of 60 months and the median is 68 months, there could still be dozens of sentences with minimums shorter than 60 months.) Figures 5a–5d show us the specific cells in which at least half the sentences were downward departures.

The figures also indicate that the size of the departures was substantial. For instance, in the murder group, the medians were from 30 to 59 months below the low end of the indicated ranges. While the departures in the other offense groups were not all equally dramatic, in the assault group, the medians were from two to 65 months below the min-mins; in the CSC group, the medians were from 1.5 to 57 months below the min-min; in the robbery group, the medians were from 1.5 to 55.5 months below the min-mins. That is, every group had cells in which judges felt the recommended minimum was at least four and a half years too long for half the sentences that scored there.

Clearly, the extent to which there were medians below the min-min varied greatly among the offense types. On the murder grid, this occurred only among the higher PRV levels, where judges apparently felt that despite the defendant's prior record the guidelines recommendations were excessive. On the CSC grid, these below-range medians were rare but were also confined to PRV levels E and F. The assault grid was quite different, with below-range medians at five of the six PRV levels and five of the six OV levels. On the robbery grid, the numerous cells with below-range medians were heavily concentrated at OV levels IV–VI.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Another sort of departure, not sanctioned by law and not visible from the grids, should be noted here. MCL 750.520b(2)(b) was amended in 2006 to require a mandatory minimum sentence of 25 years for convictions of first-degree criminal sexual conduct if the defendant was 17 or older and the victim was younger than 13.[1] The statute contains no provision for judges to depart below 25 years. The sentences on our CSC grid include 110 that were eligible for this mandatory minimum and thus should have been at least 25 years. However, 80 of them, or 72%, were less than 25 years. In fact, they ranged from one to 23.75 years with a median of 10.8 years. As the following distribution shows, half of them were below the midpoint of their applicable guidelines range.

- Below min-min – 17.5%
- At min-min – 11.3%
- Below midpoint – 22.5%
- At midpoint – 7.5%
- Above midpoint – 18.8%
- At max-min – 16.3%
- Above max-min – 6.3%

Although the numbers are not large, there is a statistically significant correlation between departing from the mandatory minimum and the county of conviction.[2] Among the counties with the most sentences eligible for the mandatory minimum, Kent departed in 16 of 17 sentences (94%), Wayne departed in 12 of 15 sentences (80%), and Washtenaw and Oakland departed in 100% of their eligible sentences (10 and nine, respectively), while Macomb departed in only 64% (nine of 14 eligible sentences), and Saginaw departed in just 33% (four of 12 eligible sentences).

We do not know how many of these CSC sentences were the result of plea negotiations that focused on avoiding the mandatory minimum while still obtaining a conviction of the offense that required it. We do know that 15% resulted from jury trials. In every case, two things are certain: judges felt the mandatory 25-year minimum sentence was inappropriately high, and prosecutors either agreed or at least did not disagree enough to have the lower sentences overturned on appeal.

C. Medians at the Minimum-Minimum

Also noteworthy are the many median sentences that were right at the low end of the applicable guidelines range. These medians represent the lowest sentences that could be imposed without requiring justification for downward departures. The CSC group had none of these medians, and the murder group had only one, in the F-II cell. However the other two offense groups tell a very different story.

The assault group had eight cells with medians that equaled the min-min. As with the downward-departing medians, the ones at the min-min were at nearly every OV and PRV level. When the two sets of medians are combined, we find that 17—nearly half of the cells on the assault grid—had medians at or below the bottom of the recommended guidelines range.

The robbery group was even more extreme. It had 19 cells with medians that equaled the min-min. Most of them were at offense levels I–III. When the medians at and below the min-min are combined, they accounted for 30 of the 36 cells on the grid, or 83%.

The extent of medians at and below the low end of the guidelines ranges does more than show a general judicial predisposition to impose lower sentences than the guidelines allow, particularly for robbery and assault. It raises the question of what the medians would look like if the starting point for the guidelines ranges against which judges measure their sentences were lower.[3]

RECEIVED by MSC 4/12/2022 1:07:49 PM

D. Same Ranges, Different Medians: Outcomes in Redundant Cells

As we observed above in Section A, the grids contain many redundant cells. The apparent goal of repeating the same ranges for different combinations of PRV and OV levels was to balance these factors in a way that achieved equivalent outcomes. To see how well this worked, pick any grid and follow the diagonal lines from the low PRV/high OV ranges on the left upward to high PRV/low OV ranges on the right. Over and over, you will find that the very same guidelines ranges produce very different median sentences.

Take, for instance, the series of cells on the CSC grid with a low end of 108 months. Moving along the six cells on the diagonal, this range started in the A-VI cell with a median of 120 months, went up to 139.5 months, moved down to 132 months, moved down further to 120 months, then jumped up to 150 months, and ended in the F-I cell, with a median of 168 months. That is, the identical ranges produced median minimum sentences that differed by as much as four years.

Consider also the series of cells on the assault grid with a low end of 135 months. Moving along the four cells on the diagonal, this range started in the C-VI cell with a median of 180 months. That dropped precipitously to a median of 135 months, declined further to 120 months, and then rose in the F-III cell by 80 months to a median of 200 months. In this instance, the identical ranges produced median minimum sentences that differed by nearly seven years in cells that were just one step apart.

The patterns among redundant cells differ depending on the offense and the guidelines range. In some instances, the high OV, low PRV cell has the highest median; in others, it is the low OV, high PRV cell. Thus, it cannot be said that judges consistently weigh either variable more heavily. But the result is largely the same. As NCSC also found, over and over again, identical ranges produce widely different median sentences.[4]

E. Different Ranges, Identical Medians: Outcomes in Non-Redundant Cells

While many redundant cells produced dissimilar results, the grids also provide examples of the opposite: cells with different ranges that had identical medians. On the murder grid, for instance, the three OV levels for PRV B had cell ranges that began with 144 months, 162 months, and 180 months. But every range had a median sentence of 216 months. Similarly, three of the six PRV levels that intersected with OV Level II—C-II, D-II, and E-II— had medians of 240 months, although the low ends of their respective ranges were 180, 225, and 270 months.

The same phenomenon is visible for the A grid offenses. For instance, for the assault group the D-IV, E-IV, and F-IV cells all had a median sentence of 120 months, even though the low ends of their ranges are 126, 135, and 171 months, respectively. That is, even though the min-mins differed by nearly four years, the medians were the same.

In the CSC group, the cells with identical medians were far from adjacent. The F-I, E-II, and F-V cells all had medians of 168 months, although the low ends of their ranges are 108, 126, and 225 months, respectively. That is, sentences in cells with starting points that differed by nearly 10 years somehow had identical medians.

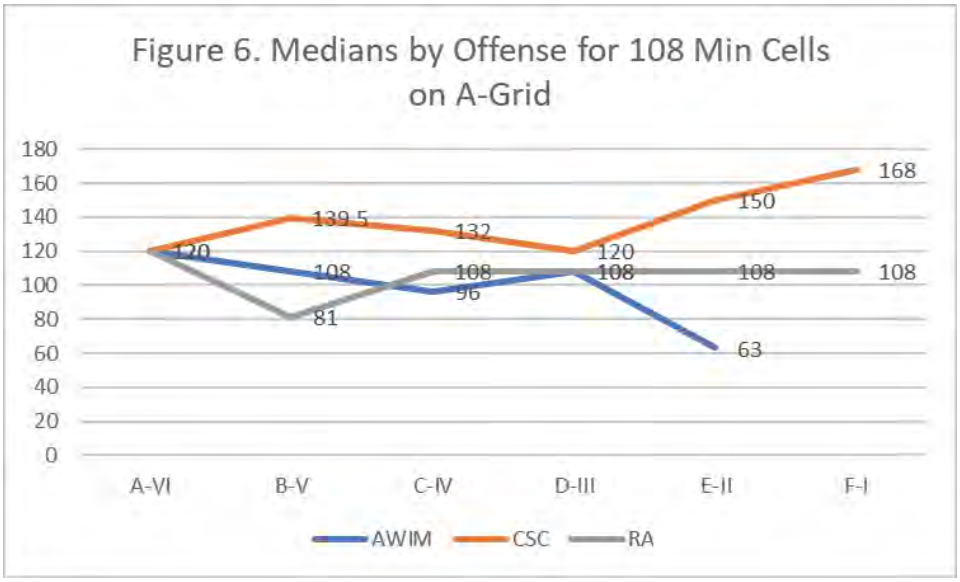
Even the robbery group shows this sort of disparity. Compare the F-IV and C-VI cells, which both had median sentences of 156 months, although the former had a starting point of 171 months and the latter's range began at 135 months. Clearly, the fact that OV and PRV scores place sentences in different ranges is no guarantee that the defendants will receive different sentences.

F. Same Ranges, Different Offenses: Medians All Over the Map

Finally, we can compare the same the cells on the A grid for different offenses. The cells with a starting point of 108 months are arrayed on the diagonal that runs from A-VI in the lower left corner of the grid to F-I in the upper right corner. Figure 6 shows where the medians fall in these cells for the assault, CSC,

RECEIVED by MSC 4/12/2022 1:07:49 PM

and robbery groups. While each offense begins with a median of 120 months, they exhibit very different patterns as they move toward the F-I cell. The CSC medians rise, the assault medians fall, and the robbery medians stay the same, resulting in a difference of more than seven years between median sentences in the very same cell.



OMNI Data, AWIM, & CSC, $p < .05$, RA $p < .001$

The jumble of outcomes we have seen throughout Chapter Eight is difficult to explain. Perhaps it simply stems from the breadth of discretion judges can exercise within the guidelines ranges combined with the high rate of both upward and downward departures. The enormous variety of choices available in every individual case inevitably undermines even the most complex effort to rationally balance the key guidelines components of offense type, criminal history, and offense severity.

RECEIVED by MSC 4/12/2022 1:07:49 PM

SUMMARY OF OUTCOMES RESULTING FROM THE GUIDELINES STRUCTURE

In adopting the Sentencing Commission's guidelines, the Legislature made numerous critical choices.

- It opted to group offenses on grids by their statutory maximum sentences instead of by the similarity of the defendant's conduct, although judges and prosecutors generally think about crimes by the nature and extent of the harm they cause.
- It included numerous redundant cells in an attempt to balance the weight given to prior record and offense severity scores.
- It made the ranges on the grids that apply to life-maximum offenses extremely broad in order to preserve judicial discretion.
- It designed prior record and offense severity variables that double-count aspects of the conviction offense and the defendant's criminal history, that score behavior that was not part of the conviction offense, and that purposefully enhance sentence lengths in murder, assault, and CSC cases.
- It incorporated numerous aggravating factors into guidelines variables, then required judges to award the maximum number of points available for those factors.
- It declined to incorporate mitigating factors into the variables.
- It extended the high end of the guidelines ranges for second, third, and fourth habitual offender sentences by 25, 50, and 100%.
- It made compliance with the guidelines mandatory but permitted departures for "substantial and compelling reasons."

It was expected that the Commission would have the opportunity to assess the impact of these choices and recommend changes as needed to better achieve the guidelines' goals. Because the Commission was abolished, that assessment never occurred. While the analyses in these chapters are no substitute for a systematic review by a multi-stakeholder, professionally staffed commission, the evidence we have examined suggests several conclusions.

- The guidelines do not compel extremely long sentences but they allow and encourage them.
    - Judges routinely imposed minimum sentences at or below the low end of the ranges, especially in armed robbery cases, even when the defendant had received a habitual offender enhancement. Judges and prosecutors often behaved as if they felt the guidelines recommendations were inappropriately harsh.
    - However, judges also commonly imposed minimum terms of 25+ years, especially on those convicted of murder, sex offenses, and as habitual offenders, while staying in compliance with the guidelines.
- The extent of departures from the recommended ranges suggests that judges were frequently dissatisfied with the options they were given.
- The complexity of the guidelines creates the appearance of precision while preserving enormous discretion for judges and enhancing the ability of prosecutors to achieve the outcomes they desire.
- The guidelines not only fail to reduce disparity, they promote it.

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER NINE. SYSTEMIC FACTORS: METHOD OF CONVICTION

The guidelines account for all the factors relevant to punishment by addressing every aspect of the defendant's criminal conduct, past and current, in exhaustive detail. Yet the regression analyses have already shown us that, because the broad guidelines ranges were designed to preserve substantial judicial discretion in life-maximum cases, and because judges nonetheless regularly depart from their recommendations, the guidelines variables do not fully account for the length of sentences imposed. This leads us to consider other factors that may affect sentence length.

These other factors can be divided into two types. First are those inherent in criminal justice case processing. These systemic variables include the method of conviction (i.e., whether the defendant pleaded guilty or went to trial), the county of conviction and the identity of the sentencing judge. Although none of these variables is logically related to the extent of punishment that is warranted by the defendant's conduct, all three have been long-recognized by practitioners and identified by prior researchers as contributing to sentence length. We will examine these factors next. A fourth systemic factor that may be related to sentence length is whether defense counsel was appointed or retained. That could not be examined because the available data did not include that information.

The second group of factors includes characteristics of the defendant not scored by the guidelines, such as age, sex, and race. We will examine these factors in Chapters Twelve and Thirteen.

The research confirms what every criminal practitioner knows: people who plead guilty get shorter sentences than people who go to trial. Indeed, sentence reductions are the primary method of inducing defendants to plead guilty, thereby sparing the emotional costs of a trial for victims, the financial costs for taxpayers, and the work for overloaded courts and prosecutors. It is a truism that the criminal justice system would grind to a halt if the majority of cases were not disposed of by negotiated pleas. The U.S. Supreme Court recognized this when it upheld the constitutionality of bargaining fifty years ago in *Brady v. United States*.[1] Yet, because the right to a jury trial is guaranteed by the United States and Michigan constitutions, the flip side of a reward for pleading guilty is a tax on the decision to go to trial.

Some argue that a bargained sentence is a benefit as opposed to a punishment for going to trial. This reasoning assumes that sentences imposed after trial are what defendants deserved and by pleading they get less than they deserve. This view further assumes that there is a "regular" sentence that clearly identifies what each defendant deserves. However, where judges have very broad discretion as in the life-maximum cases, there is no regular sentence. In fact, where plea bargains are routine, the bargained sentences become the norm. The sentences imposed at trial are an upward deviation from that norm. And the only thing that distinguishes the cases may be the fact of a trial. That fact is unrelated to any of the traditional goals of sentencing: retribution, deterrence, incapacitation, rehabilitation. On occasion, a particularly harsh sentence after trial can be explained by information the judge has gained about the offense or the defendant (e.g., through witness testimony or the defendant's behavior in court). However this is not sufficiently common to account for the routine differential in sentences after trial.[2] The bottom line is: "Whether one calls this a punishment for going to trial or the failure to receive a benefit for pleading guilty, the fact that differential sentencing occurs is indisputable."[3]

A. The Risks of Plea Bargaining and the Decision to go to Trial

Concern about the coercive effect of plea bargaining on the waiver of the constitutional right to jury trial is more than theoretical. The critical practical risk is that a defendant who is factually innocent or who would, at least, have a substantial chance of acquittal at trial, will be induced by a highly generous offer or the threat of extremely harsh punishment to forego the opportunity to have his or her guilt determined by a jury.

The decision is by far the most complex in life-maximum cases. Obviously, the risk of a very lengthy sentence is the greatest in these cases. Consequently the prosecution has far more leverage. The inducements may be dismissed charges, an agreement not to bring charges, or the promise, express or implied, of a lower sentence. The multiple terms of the bargain in *People v. Biggs* illustrates this point.

RECEIVED by MSC 4/12/2022 1:07:49 PM

The defendant was charged with the life-maximum offenses of armed robbery and assault with intent to murder, plus felony firearm, which carries a mandatory consecutive sentence of two years. The agreement was that the defendant would plead guilty to the robbery in exchange for dismissal of the assault and firearm charges, a promise that the prosecution would not file a habitual offender enhancement and a sentence recommendation of 12 to 25 years.[4]

Plea bargaining does not only affect the guilty. It creates pressure on defendants who might well be acquitted or convicted of a lesser offense to forego the chance to have a jury weigh the credibility of prosecution witnesses or consider evidence of defenses such as alibi, mistaken identification, self-defense, provocation, or lack of intent.[5]

On the other hand, life-maximum cases are also the ones in which prosecutors who have strong evidence may have the least incentive to make an attractive offer or to make an offer at all. The seriousness of the crimes, press attention, or objections by victims may all lead prosecutors to refuse to negotiate, giving defendants little or no incentive to plead. Ironically, just as extreme pressure may induce even innocent defendants to waive their constitutional right to a trial, the fact that there is no right to a beneficial bargain means that guilty defendants who would be willing to accept a reasonable offer may be forced to go to trial.

A note about bench trials

A trial just before the court, without a jury, is often referred to as a bench trial. It is a simpler, more efficient proceeding than a jury trial. Logistically, there are fewer people to assemble and accommodate. There is no need for jury selection or jury instructions, and ultimately only one person must be persuaded of the defendant's guilt.

Bench trials are sometimes chosen for strategic reasons related to the complexity or nature of the evidence. If a case is highly technical or highly emotional, it may be preferable for the fact-finder to be a professional with experience in such cases.

However, a bench trial may also be a compromise between pleading guilty and an expensive, time-consuming jury trial. The defendant with little chance of persuading a jury to acquit nonetheless gets the chance to present a defense, while losing the chance to hang a jury and require a retrial. The hope is that a sympathetic judge may convict of a lesser offense or at least reward the waiver of a jury with a sentence reduction. As the data show, the trial tax is generally lower when the trial is by the court. For this reason, bench trials are sometimes referred to as "slow pleas of guilty."

In most jurisdictions, bench trials are rare. In the former Recorder's Court for the City of Detroit, they were relatively common. Although their numbers declined after Recorder's Court was abolished and its judges joined the Wayne County Circuit bench, the proportion of bench trials in Wayne County still far exceeded that of other counties in both of our study periods.

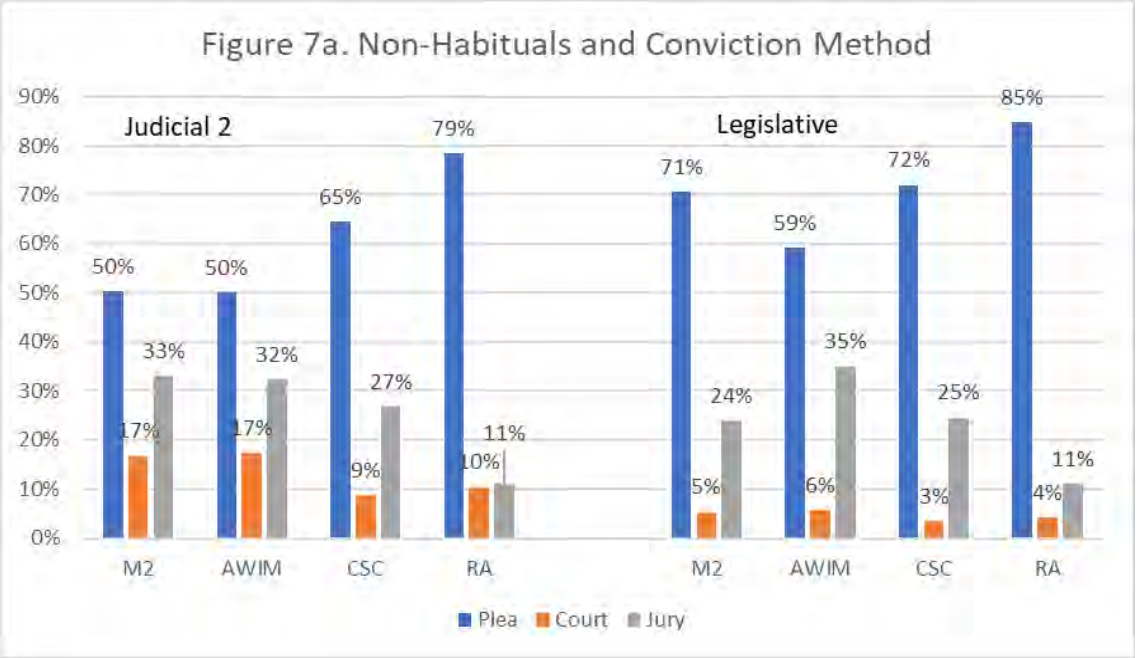
B. Frequency of Conviction Methods

The frequency of dispositions by plea has long been less for life-maximum offenses than for all other crimes. The State Court Administrative Office (SCAO) keeps separate statistics for "capital" (i.e., life-maximum) offenses. In 2019, 97.4% of non-life-maximum cases were resolved by guilty plea. The comparable figure for life-maximum cases was 79.6%,[6] a proportion that was actually skewed by the large number of armed robbery cases that are more likely than other crimes to result in a plea.

Figure 7a shows the methods of conviction by offense type for non-habitual offender sentences in the second judicial and legislative guidelines periods. Note that under both sets of guidelines, the rate of pleas varied by offense type but increased markedly for each offense under the legislative guidelines. For all but murder sentences, the increase in pleas resulted from a decline in bench trials while jury trial rates stayed roughly constant. Among murder sentences, plea rates exploded while the proportion of both bench and jury trials declined. The greater change for murder sentences may be related to the fact that many involve a plea down from first-degree murder with its mandatory sentence of life without parole.[7]

RECEIVED by MSC 4/12/2022 1:07:49 PM

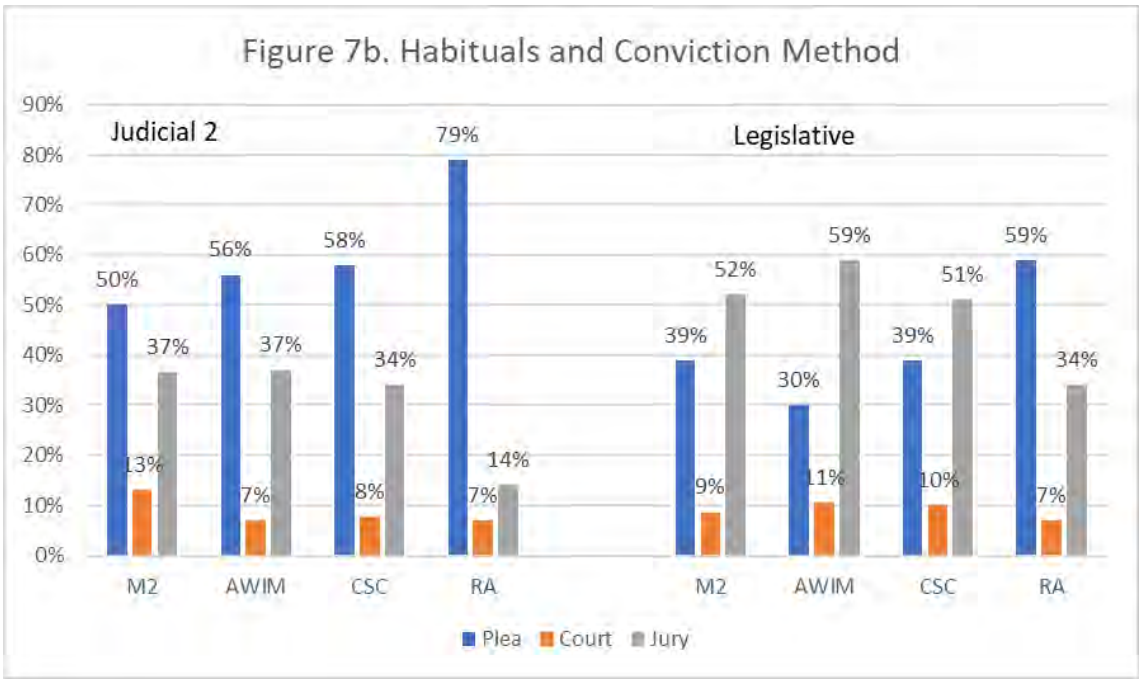
Despite the increase in pleas after 1998, the plea rate for non-habitualized defendants in life-maximum cases remained much lower than that in non-life-maximum cases. The numbers alone cannot tell us which of the two opposite explanations are at work—whether prosecutors make fewer substantial offers in life-maximum cases, or defendants insist on going to trial because they believe they have viable defenses.



CMIS data, P < .001 CMIS data, P <.001

Figure 7b shows the conviction methods for habitual offender sentences in each offense group during each time period (i.e., those of defendants who did not have habitual enhancements dismissed or who, at most, had them reduced). After the change from judicial to legislative guidelines, when habitual offenders were covered by guidelines for the first time, the proportion of pleas dropped dramatically and the proportion of jury trials increased in an equally striking fashion. For every group except murder, pleas declined and jury trials increased by roughly 20 points. Murder cases showed the same trend to a lesser degree. The differences cannot be explained by a shift in the use of bench trials. For one thing, this shift was very small. Moreover, when bench trials declined for non-habituals, the proportion of pleas increased.[8]

RECEIVED by MSC 4/12/2022 1:07:49 PM



CMIS data, P < .001 CMIS data, P <.001

When we compare Figures 7a and 7b, we can see how much conviction methods underlying habitual and non-habitual sentences differed from each other in each time period. When the second judicial guidelines, which did not cover habitual sentences, were in effect, jury trial rates for habitual and non- habitual sentences were quite similar. For each offense group, the jury trial rate for the habitual group was three to seven points higher. The difference in plea rates was zero to seven points lower for the habitual group.

Under the legislative guidelines, the differences in conviction methods between habitual and non-habitual sentences became stark. As habitual enhancements led to far fewer pleas for every offense type, the overall rate of jury trials nearly doubled, from 21.7% during the judicial period to 40.8% under the legislative guidelines.

The data show us only which sentences were and were not habitualized, not any underlying plea negotiations. We cannot see how many habitual enhancements were brought and dismissed in each time period or how many enhancements were threatened but not brought as part of a plea bargain. However, if prosecutors used the threat of having the minimum sentence increased by 25, 50, or 100% primarily to create leverage in plea negotiations, one would expect the rate of pleas to rise, not fall. If prosecutors refused to reduce or eliminate habitual enhancements because their priority was increasing sentence lengths, that would explain the decline in habitual pleas under the statutory guidelines.

C. Relationship to Sentence Length: The Size of the Trial Tax

Whether defendants go to trial because they want to assert defenses or simply for lack of better options, they face the prospect of receiving a longer prison term for having done so. This result raises questions the data cannot answer.

- Are sentences longer after a trial simply because the defendants lack the benefits of a negotiated plea?
- Is it because the court had the opportunity at trial to consider particular evidence or the demeanor of the defendant or that of prosecution witnesses?

RECEIVED by MSC 4/12/2022 1:07:49 PM

- Or is it a deliberate penalty for the defendant's decision to "waste" the court's time by insisting on exercising his or her Sixth Amendment right?
- Even more fundamentally, what is the relationship of the conviction method to the purposes of sentencing? If 15 years is adequate to punish the defendant and protect the public when someone pleads guilty, why is 25 years necessary when the same person goes to trial?

Whatever the answers, the trial tax existed before the adoption of judicial guidelines and continued to exist under the legislative guidelines. The trial tax is calculated as the percentage increase in sentence length for jury trials over pleas. Thus, for instance, if the median sentence for defendants who pleaded guilty was 10 years, and for people convicted by a jury it was 20 years, the trial tax would be 100%.

Table 14 summarizes the median sentences and the extent of the trial tax for both non-habitual and habitual sentences during the judicial period. While the habitual sentences for pleas were much higher than the non-habitual sentences for pleas, the actual number of years added for going to trial was not. The number of years non-habitualized defendants paid for going to a jury trial was between five and seven. The number of years habitualized defendants paid was between five and eight.

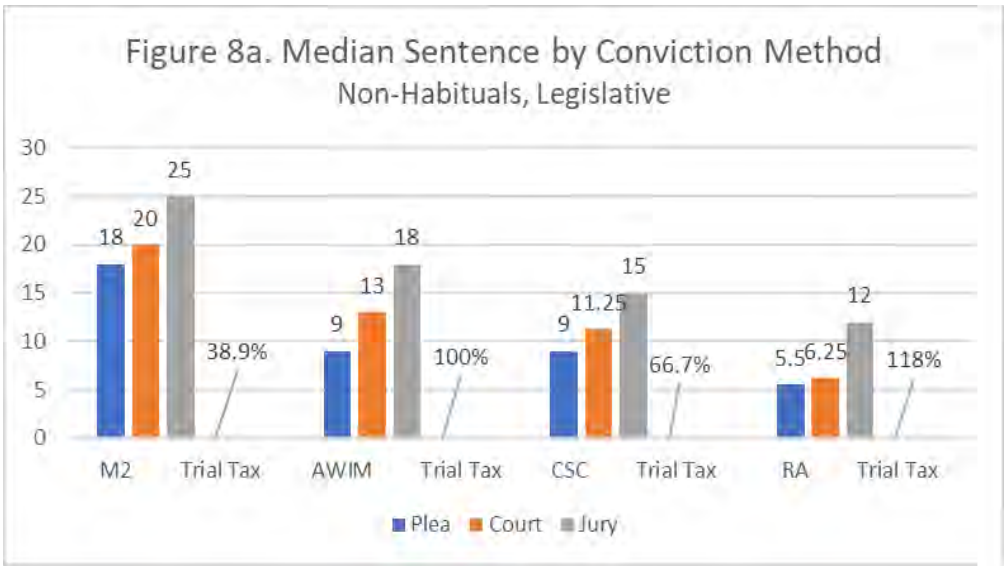
Table 14. Median Sentences by Conviction Method During 2nd Judicial Guidelines Period Non-Habitual and Habitual Sentences (in years), 1989–1998

	Non-Habitual			Habitual		
Crime	Plea	Jury Trial	Trial Tax	Plea	Jury Trial	Trial Tax
Murder, second-degree	15.0	20.0	5 yrs/33.3%	25.0	30.0	5 yrs/20.0%
Assault	8.0	15.0	7 yrs/87.5%	12.0	20.0	8 yrs/66.7%
CSC, first-degree	8.0	15.0	7 yrs/87.5%	17.25	23.75	6.5 yrs/33.7%
Robbery	4.0	10.0	6 yrs/150.0%	7.5	15.0	7.5 yrs/100.0%
CMIS Data	All offense types significant at $p < .001$			M2 $p < .05$, AWIM, CSC, & RA $p < .001$		

Figure 8a illustrates the pattern of sentence increases from plea to bench trial to jury trial for non-habitual sentences under the legislative guidelines. The baseline was increased compared to the judicial period in that sentences for pleas of every offense type were longer. Except for CSC, sentences for jury trials were longer too.

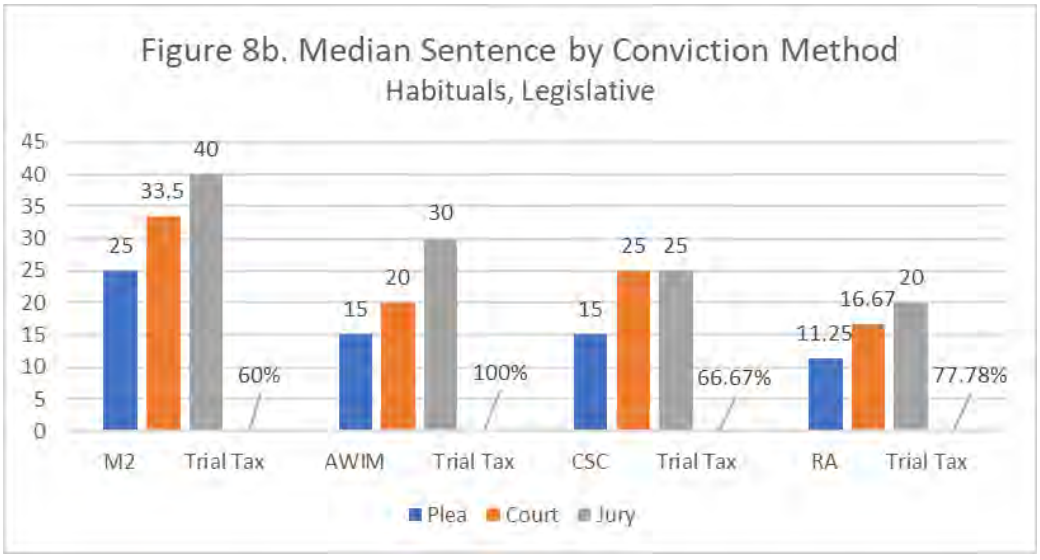
For every offense type the sentences for bench-tried defendants were between those for guilty pleas and jury trials but closer to pleas. The number of years defendants paid for going to a jury trial was between six and nine.

RECEIVED by MSC 4/12/2022 1:07:49 PM



CMIS Data, All offense types significant at $p < .001$

Figure 8b indicates that, compared to the judicial period shown in Table 14, under the legislative guidelines the trial tax on habitual offender sentences exploded. The percentage increase for jury trials over pleas was much greater for every offense group except robbery. The actual years added for requesting a jury trial ranged from nearly nine in robbery cases to 15 in murder and assault cases. Even having a bench trial cost habitualized defendants an extra five to10 years in prison.



CMIS Data, All offense types significant at $p < .001$

Thus defendants sentenced as habitual offenders for life-maximum offenses under the legislative guidelines pay three times over. The guidelines ranges for these offenses are higher for all defendants than they were under the judicial guidelines. The ranges are then broadened by enhancement as a habitual offender. And habitual offenders are far more likely to be convicted at a jury trial, which exposes them to a substantial trial tax.

The trial tax reflects not only the fact that the defendant went to trial but all the factors associated with that decision, including the defendant's prior record, the details of the offense, habitual offender

RECEIVED by MSC 4/12/2022 1:07:49 PM

status, and the likelihood of a departure. When the regression analysis honed in solely on the impact of the conviction method while holding all other variables constant, it appeared that:

- Just the fact of having a jury trial added 18% to the sentence compared to a guilty plea, while the impact of a bench trial increased sentence length by 6%. These results are similar to NCSC's finding that conviction at trial increased the sentence by 17.7%.
- The amount of the increase varied by offense—ranging from a 12% increase for murder cases to a 21% increase when the conviction was for assault.
- As we will see in Chapter Ten, the sentence increases for selecting a jury trial varied substantially by county.

D. How the Trial Tax Is Applied Under Sentencing Guidelines

The medians show how greatly sentences differ depending on whether the defendant was convicted by guilty plea or jury trial. But how, exactly, do the sentencing guidelines allow these wide disparities to occur? Our discussion in Chapter Seven has already suggested the answers: the breadth of the cell ranges and the amount of departures above and below those ranges.

Together, these factors create so much room for discretion that a non-guidelines consideration like conviction method can have a striking impact on the sentence. Table 15 shows that, for every type of offense, those who pleaded guilty were far more likely to receive a sentence below the bottom of the applicable cell range while those convicted by a jury were far more likely to receive a sentence right at or above the highest end of the range. Overall, plea-based sentences were more than twice as likely to be below the midpoint of the applicable range.

Table 15. Percentages of Minimum Non-Habitual Sentences across Guideline Ranges, by Conviction Method, 2003–2012

	Number of Cases	Below Min-Min	At Min-Min	In Range, Below Midpoint	At Midpoint	In Range, Above Midpoint	At Max-Min	Above Max-Min
M2	**1,218**							
Plea	896	39.8	5.9	19.4	9.5	10.2	5.1	10.0
Court	49	22.4	12.2	22.4	10.2	16.3	10.2	6.1
Jury	273	9.2	5.1	15.0	17.9	17.9	16.8	17.9
AWIM	**884**							
Plea	553	55.0	9.6	16.6	2.9	7.8	1.6	6.5
Court	38	21.1	13.2	13.2	5.3	21.1	2.6	23.7
Jury	293	9.9	7.5	15.0	10.2	25.6	14.7	17.1
CSC	**1,225**							
Plea	873	28.3	12.4	21.0	7.4	12.5	9.4	9.0
Court	38	23.7	0.0	15.8	5.3	18.4	5.3	31.6
Jury	314	9.2	5.1	15.0	10.2	18.8	17.8	23.9
RA	**5,659**							
Plea	4,915	43.3	19.4	19.1	4.5	7.6	1.8	4.3
Court	210	37.1	22.9	13.8	6.2	8.1	4.3	7.6
Jury	534	13.9	10.9	19.3	9.2	17.6	14.4	14.8
OMNI Data, All offense types significant at $p < .001$								

RECEIVED by MSC 4/12/2022 1:07:49 PM

To test the possibility that this distribution of sentences may reflect different combinations of offense and prior record scores (e.g., that downward departures involve less serious cases), we looked at the distribution of sentences within the range most commonly used for each offense. The patterns were similar to those in Table 15. That is, people whose OV and PRV scores placed them in cells with identical parameters received very different sentences depending on whether they pleaded guilty or went to trial.[9]

Among habitual offender sentences the differences in departures above the usual maximum-minimum between pleas and jury trials were even more extreme. For example:

- When the defendants were sentenced for murder as second offenders, 35% of the plea-based sentences were downward departures and only 8.8% were above the max-min. Of those convicted by a jury, 2.9% of the sentences were downward departures and 57% were above the max-min.
- Among second-offense assault sentences, 32% of the pleas but none of the jury trials resulted in downward departures. Conversely, 12% of the pleas and 76% of the jury trials resulted in sentences above the max-min.
- Assault defendants treated as fourth offenders who were jury-tried received downward departures in just 2% of their sentences and terms above the max-min in 75%. By comparison, among plea-based sentences 31% received downward departures while 21% had terms above the max-min.
- Fourth offender CSC sentences reflected downward departures in 18% of the pleas and none of the jury trials. Conversely, 30% of the plea-based sentences and 66% of the jury-based sentences were above the max-min.
- Fourth offense armed robbery sentences showed the same pattern. Only 1% of the jury-based sentences were downward departures compared to 19% of the pleas. At the other extreme, nearly 60% of the jury-based sentences but only 18% of the plea-based sentences were above the max-min.

In sum, sentences for every type of offense that resulted from guilty pleas were far more likely to be at or below the low end of the applicable guidelines range and far more likely to be at or above the high end if they resulted from jury convictions. This would not have a major impact on sentence length if the ends were not so far apart. However, given the breadth of the guidelines ranges, the use of habitual offender charges to broaden them even further and the routine use of both upward and downward departures, it becomes easy to see how the price of a jury trial may be years or even decades of additional incarceration.

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER TEN. SYSTEMIC FACTORS: COUNTY OF CONVICTION

The current analysis confirms what previous research has shown about the impact of locality. Whether Michigan has no guidelines, advisory judicial guidelines, or mandatory legislative guidelines, defendants with similar backgrounds who commit similar offenses may receive very different sentences depending on the county of conviction. To assess the differences in sentence lengths among counties, we must first recognize that there are significant differences in the frequency of offenses. A disproportionate share of offenses that typically carry longer sentences could explain why some counties have longer median sentences. This possibility reinforces the importance of controlling for offense type in analyzing the data.

Table 16. Frequency of Offense Type by County, Non-Habitual Sentences, 1999–2012

	Murder	Assault	CSC	Robbery	Total
Berrien	21	14	43	139	217
	9.7%	6.5%	19.8%	64.1%	
Calhoun	15	20	74	74	183
	8.2%	10.9%	40.4%	40.4%	
Genesee	144	77	70	402	693
	20.8%	11.1%	10.1%	58.0%	
Ingham	42	23	62	153	280
	15.0%	8.2%	22.1%	54.6%	
Jackson	7	6	29	73	115
	6.1%	5.2%	25.2%	63.5%	
Kalamazoo	33	23	67	182	305
	10.8%	7.5%	22.0%	59.7%	
Kent	83	67	246	771	1,167
	7.1%	5.7%	21.1%	66.1%	
Macomb	46	92	143	665	946
	4.9%	9.7%	15.1%	70.3%	
Muskegon	18	11	81	141	251
	7.2%	4.4%	32.3%	56.2%	
Oakland	59	113	267	915	1,354
	4.4%	8.3%	19.7%	67.6%	
Saginaw	39	61	68	217	385
	10.1%	15.8%	17.7%	56.4%	
Washtenaw	34	26	60	269	389
	8.7%	6.7%	15.4%	69.2%	
Wayne	1,182	865	766	4,882	7,695
	15.4%	11.2%	10.0%	63.4%	
Total	1,723	1,398	1,976	8,883	13,980
	12.3%	10.0%	14.1%	63.5%	
CMIS Data, $p < .001$					

As Table 16 shows, in most counties for non-habitual sentences, robbery is far and away the most common of the four offense types. Robbery sentences, which are generally the shortest, are at least 55% in 12 of 13 counties; they are over two-thirds of the total in Oakland, Macomb, Kent, and Washtenaw. In 10 counties, CSC is the second most frequent offense type. In Wayne and Genesee, murder sentences, which are generally the longest, are the second most common, while in Oakland, Macomb, Saginaw, and Calhoun they are fourth.

RECEIVED by MSC 4/12/2022 1:07:49 PM

A. Median Sentence

Whether the sentences in a particular county are harsher or more lenient than those of other counties varies with the offense. No single county always gives the longest or the shortest minimum terms. The extent of disparity also varies with the offense. However, as Table 17 shows, there are definite patterns among the non-habitual sentences.

- County size appears to be more important than location.[1]
  - Sentences in the largest counties were more similar to each other than they were to the sentences imposed by mid-sized counties that are closer by. So, for instance, Kent sentences were generally more like those in Macomb, Oakland, and Wayne than they were like other counties on the west side of the state. Genesee, the fifth largest county, did not cluster with the other large counties as consistently.
  - Differences among the medium and smaller counties do not appear to depend on geography. Counties that are right next to each other, like Kalamazoo and Calhoun, had very different median sentences for CSC and robbery.
- The median sentences imposed in the larger counties tended to be shorter than those imposed in smaller counties.
- The data undermine some widely held assumptions about which counties are "tough" and which are "lenient." Wayne County, home to the City of Detroit, did not impose the lowest median sentences in any offense group. Oakland County sentences were in the low-to-middle range for every offense, despite the county's "law and order" reputation.
- Despite the fact that the M2 grid has the widest cell ranges by far, murder sentences showed the greatest amount of consistency, with nine counties at 20 years and four from 22.5 to 24 years.
- Assault sentences trended markedly higher than CSC sentences.
- CSC showed the least consistency among counties, with a spread of nearly nine years between the lowest medians and the highest.
- Although armed robbery had much lower sentences than the other offenses, the four-year spread between the lowest and highest counties was the same as the spread for murder.

Table 17. Median Sentence by County, 1999–2012 (non-habitual)

Murder, Second-Degree			Assault to Murder			CSC, First-Degree			Armed Robbery		
County	Yrs	N	County	Yrs	N	County	Yrs	N	County	Yrs	N
Berrien	20	21	Kent	10	67	Kalamazoo	6.75	67	Macomb	5	665
Calhoun	20	15	Macomb	10	92	Oakland	9	267	Oakland	5	915
Genesee	20	144	Oakland	10.5	113	Washtenaw	9.5	60	Kalamazoo	6	182
Jackson	20	7	Wayne	11	865	Kent	10	246	Kent	6	771
Kalamazoo	20	33	Jackson	11.25	6	Macomb	10	143	Wayne	6	4,882
Kent	20	83	Muskegon	12	11	Wayne	10	773	Muskegon	6.5	141
Macomb	20	46	Genesee	14	77	Calhoun	10.6	74	Saginaw	6.75	217
Oakland	20	59	Ingham	14	23	Berrien	11.25	43	Washtenaw	6.75	269
Wayne	20	1,182	Washtenaw	14	26	Muskegon	12	81	Ingham	7	153
Saginaw	22.5	39	Kalamazoo	15	23	Genesee	13.3	70	Jackson	7	73
Ingham	22.7	42	Saginaw	15	61	Saginaw	13.5	68	Genesee	7.1	402
Washtenaw	23	34	Calhoun	15.5	20	Ingham	14.5	62	Berrien	8	139
Muskegon	24	18	Berrien	16.25	14	Jackson	15	29	Calhoun	9	74
Spread	4		Spread	6.25		Spread	8.75		Spread	4	
CMIS data, Not Sig			p < .05			p < .001			p < .001		

RECEIVED by MSC 4/12/2022 1:07:49 PM

Regression analysis confirms these findings. Examining estimated sentence lengths for all offenses combined and using Wayne County as the baseline found that:

- The length of sentence overall was significantly longer in seven counties compared to Wayne, significantly shorter in three, and not much different in the remaining two.
    - For instance, Genesee, Jackson, and Muskegon had sentences that were 7% longer while Oakland's were 7% shorter.
    - Similarly, Kent's sentences were 4% longer than Wayne's while Macomb's were 4% shorter.
- The disparities by county were far and away most pronounced for CSC sentences.
    - Eight counties had sentences that were significantly different than Wayne's and all were longer.
    - The greatest difference was with Genesee, which was 24% longer; the smallest was Kent which is 5% longer.

B. Place on the Range

The difference in median sentence length is related to how a county's judges utilize the often very broad guidelines ranges. In Chapter Seven we explored how "place on the range," that is, the extent to which sentences are at the high or low end of the applicable cell range, or depart from the range altogether, varied depending on the nature of the offense and the defendant's habitual offender status. Table 18 allows us to examine how place on the range varies by county without regard to offense type. By rank ordering the counties at points below, within, and above the range we can more readily see patterns. Each county's percentages of non-habitual sentences below and at the minimum-minimum are added to create a "low end" total; each county's percentages of sentences at or above the maximum-minimum are added to create a "high end" total.

The great many departures below the guidelines' recommended range cannot be explained by county size or location.

Downward departure rates vary greatly by county. For instance, rates for the four largest counties decline in increments of roughly 10 points:

- Wayne – 47.6%
- Kent – 36.0%
- Macomb – 26.4%
- Oakland – 16.5%

The medium-sized counties also differ from each other; for example, the rate of downward departures is 35.6% in Kalamazoo but only 17.2% in Washtenaw.

The smaller counties of Berrien, Muskegon, Calhoun, and Jackson show more consistency, with fewer than 10% of their sentences being downward departures.

Overall there are more than 4.5 times as many downward as upward departures.

- Most counties depart down at anywhere from four to 10 times as often as they depart upwards.
- The smaller counties of Berrien, Muskegon, Calhoun, and Jackson each departs both up and down at small but more similar rates.

Table 18. Place on the Range by County, Rank Ordered by Percentage of Non-Habitual Sentences, 2003–2012													
Below Min-Min		At Min-Min		Low-End Total		Below/At/Above Midpoint Total		At Max-Min		Above Max-Min		High-End Total	
Wayne	47.6	Washtenaw	42.4	Kalamazoo	65.4	Muskegon	71.8	Genesee	20.4	Wayne	10.9	Calhoun	25.9
Kent	36.0	Oakland	36.5	Washtenaw	59.6	Jackson	63.8	Berrien	18.3	Ingham	9.8	Genesee	25.2
Kalamazoo	35.6	Kalamazoo	29.8	Wayne	56.3	Berrien	52.3	Calhoun	17.9	Calhoun	8.0	Berrien	24.6
Ingham	32.6	Macomb	29.6	Macomb	56.0	Calhoun	50.0	Jackson	16.3	Saginaw	7.0	Jackson	22.6
Genesee	28.0	Calhoun	20.5	Oakland	53.0	Kent	47.2	Saginaw	12.0	Jackson	6.3	Saginaw	19.0
Macomb	26.4	Saginaw	18.6	Kent	45.3	Saginaw	42.1	Ingham	7.3	Berrien	6.3	Ingham	17.1
Saginaw	20.2	Berrien	14.3	Ingham	41.9	Ingham	40.9	Macomb	7.1	Muskegon	5.8	Wayne	13.4
Washtenaw	17.2	Genesee	10.5	Saginaw	38.8	Oakland	39.2	Muskegon	6.4	Kalamazoo	4.9	Muskegon	12.2
Oakland	16.5	Ingham	9.3	Genesee	38.5	Genesee	36.4	Oakland	5.3	Genesee	4.8	Macomb	11.2
Berrien	8.7	Kent	9.3	Calhoun	24.1	Washtenaw	33.6	Kent	4.2	Washtenaw	4.2	Kalamazoo	8.3
Muskegon	7.7	Jackson	8.8	Berrien	23.0	Macomb	32.8	Kalamazoo	3.4	Macomb	4.1	Oakland	7.7
Jackson	5.0	Wayne	8.7	Muskegon	16.0	Wayne	30.3	Washtenaw	2.7	Kent	3.2	Kent	7.4
Calhoun	3.6	Muskegon	8.3	Jackson	13.8	Kalamazoo	26.3	Wayne	2.5	Oakland	2.4	Washtenaw	6.9
OMNI Data, Chi-Square p < .001													

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

The obverse of departure rates is the extent to which counties comply with the sentencing guidelines. Compliance can be seen by adding the second, fourth, and fifth columns in Table 18.

- Wayne, the county with the largest number of cases by far, imposes sentences within the recommended range only 41.5% of the time. At 57.6, 59.5, and 60.8%, respectively, the counties with the next lowest compliance rates are Ingham, Kalamazoo, and Kent.
- At the other extreme, the smaller counties of Berrien, Muskegon, Calhoun, and Jackson impose within-range sentences between 85 and 89% of the time.
- Notably, Oakland at 81.1% and Washtenaw at 78.6% behave much more like the smaller counties than like their larger peers.

As we also observed in Chapter Seven, judges who do not wish to depart, with the requirement for justification that departures entail, may nonetheless gravitate toward the very lowest or very highest minimum sentences that define the boundaries of the applicable range. The extent to which judges chose sentences at the min-min, below it, or both varied, again without regard to size or location.

- While Oakland and Washtenaw judges did not often depart below the range, 42.4 and 36.5% of their sentences, respectively, were the minimum-minimum.
- Wayne, Ingham and Kent, which all had high rates of downward departures, all imposed the minimum-minimum in only about 9% of their sentences.
- Kalamazoo and Macomb each had substantial rates of both downward departures and imposition of the minimum-minimum.

When the downward departures and minimum-minimums are combined, we get a picture of the extent to which counties impose sentences at the low end.

- Five counties: 53.0 – 65.4%
- Four counties: 38.5 – 45.3%
- Four counties: 13.8 – 24.1%

To a much smaller degree, we see a similar phenomenon with judges who did not often depart upward but who imposed a substantial number of sentences at the maximum-minimum. For the most part, the counties that had the most low-end sentences had the fewest at the high end.

The proportion of low-end sentences suggests that judges in the larger and medium-sized counties saw the guidelines recommendations as too high. But we must confirm that this appearance is not an artifact of the frequency with which these judges see certain offense types. To do so, we separate the sentences by offense type, then compare each county's proportion of low-end and high-end sentences. This process allows us to see three things: (1) the extent to which the proportion of low-end and high-end sentences each vary by offense, (2) the extent of disparity among counties for sentences for the same offense, and (3) whether there are any patterns in the relationship between high- and low-end sentences.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 19a. County Differences in Low and High-End Non-Habitual Sentences, 2003–2012		
Murder, Second-Degree		
County	Percent at Low End	Percent at High End
Macomb	44.1	26.5
Genesee	43.3	29.2
Wayne	41.0	17.8
Washtenaw	38.4	7.6
Kalamazoo	33.4	36.4
Ingham	30.4	30.4
Saginaw	30.0	20.0
Jackson	25.0	50.0
Kent	23.4	20.0
Oakland	20.9	13.9
Calhoun	18.2	36.4
Berrien	10.0	30.0
Muskegon	0.0	22.0
OMNI Data, p < .001		

We begin by looking at the non-habitual sentences for murder, which tend to be very similar. Table 19a makes two things apparent. First, there are wide disparities across counties, with the proportion of low-end sentences ranging from 44% down to zero and the proportion of high-end sentences ranging from 50% down to 7.6. The differences cannot be explained by county size or location. Compare, for example, the low-end rates of Macomb and Wayne with those of Kent and Oakland. Compare as well the high end rates of Washtenaw and Oakland with those of Calhoun and Kalamazoo.

Second, there is no consistent relationship within counties between the rates of high- and low-end sentences. Some counties had substantially more low-end sentences, some had more at the high end and some had quite even proportions at both ends. Compare, for instance, Wayne and Washtenaw with Kent and Kalamazoo.

Table 19b has the non-habitual sentences for the offenses on the A grid. It shows significant disparity by county both within and across offense types.

- The proportion of low-end sentences for the same offense varies widely across counties. Compare, for example, Genesee with Macomb or Kalamazoo with Wayne.
- The proportion of high-end sentences for the same offense also varies widely across counties. Compare, for example, Berrien with Washtenaw or Oakland with Genesee.
- For robbery the rates of low-end sentences were exponentially higher than for any other offense.
    - For eight counties, more than half the robbery sentences were at or below the min-min.
    - Among the four counties with low end rates between 61 and 67%, the proportion attributable to downward departures differed dramatically with Wayne at 51%, Macomb at 30%, Oakland at 20%, and Washtenaw, at just 14.5%.
    - Kalamazoo's low-end rate of 80% was achieved by a combination of 40% downward departures and 40% at the min-min.
- Overall, CSC sentences were markedly different from the other offense types. Eight of the counties had less than 25% of their CSC sentences at the low end, while the same eight counties had nearly 25% of their CSC sentences at the high end.
- The extent to which counties tended to be consistent in their rate of low-end sentences across offense types varied.
    - Genesee, for example, had roughly 40% of its sentences for every other offense type at the low end but only 7.5% of its CSC sentences. Macomb and Ingham also saw large declines in their CSC sentences.
    - On the other hand, Washtenaw, Oakland, and Kent each had roughly similar rates of low-end sentences for assault and CSC with lower rates for murder.
    - Kalamazoo was unique in having a much greater percentage of low-end sentences for CSC than for assault or murder.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 19b. County Differences in Low and High-End Non-Habitual Sentences for Assault, CSC, and Robbery, 2003–2012								
Assault with Intent to Murder			Criminal Sexual Conduct			Armed Robbery		
County	Percent at Low End	Percent at High End	County	Percent at Low End	Percent at High End	County	Percent at Low End	Percent at High End
Macomb	53.6	12.7	Kalamazoo	52.8	16.7	Kalamazoo	79.8	7.0
Wayne	52.8	18.3	Washtenaw	49.0	19.2	Washtenaw	66.8	3.5
Washtenaw	47.0	5.9	Wayne	44.8	21.9	Macomb	64.0	7.0
Oakland	41.7	8.4	Kent	36.4	17.6	Wayne	62.8	10.0
Ingham	41.2	11.8	Oakland	35.7	15.9	Oakland	61.0	4.9
Kent	40.5	7.2	Macomb	23.6	24.7	Ingham	53.6	10.8
Genesee	40.0	26.7	Ingham	17.0	29.3	Saginaw	52.6	8.3
Kalamazoo	38.1	19.0	Calhoun	15.2	41.3	Kent	50.7	3.1
Berrien	33.3	22.2	Saginaw	13.2	50.0	Genesee	40.9	18.5
Saginaw	24.4	24.4	Genesee	7.5	55.0	Calhoun	35.5	11.1
Calhoun	20.0	10.0	Muskegon	1.9	28.3	Berrien	32.0	14.1
Muskegon	12.5	12.5	Berrien	0.0	51.7	Muskegon	26.8	1.2
Jackson	0.0	0.0	Jackson	0.0	54.5	Jackson	19.2	7.7
OMNI Data	$p < .001$			$p < .001$			$p < .001$	

C. Conviction Method

Disparities in sentence length may also be affected by differences in the rate of pleas versus jury trials. Recall that median sentences resulting from pleas are substantially shorter than sentences resulting from jury trials and are typically shorter than those resulting from bench trials as well. Recall also that Wayne County is unique in its tradition of using bench trials. This higher rate of bench trials helps explain why Wayne County's median sentences are not among the very lowest for the three A-grid offenses.

Table 20 shows that, within each offense type, the counties vary greatly in their rate of plea-based sentences. Some consistencies exist. Kent and Kalamazoo have among the highest plea rates for every offense. Genesee also has high rates while Ingham and Wayne have among the lowest for every offense except murder. Other counties have roughly similar proportions of pleas across offense types. Most have their lowest plea rates for assault and their highest for robbery.

What is not consistent is a relationship between the proportion of pleas and the proportion of low-end sentences. Given the undeniable and intentional relationship between guilty pleas and lower sentences, one would expect the counties with the highest rates of pleas to have the highest rates of low-end sentences. However, this did not prove to be the case.

The clearest example is the murder sentences. Although nine counties all had median minimums of 20 years, their rates of guilty pleas ranged from 54.3% for Macomb and 62.7% for Oakland to 85.7% for Ingham and 83.3% for Genesee.

RECEIVED by MSC 4/12/2022 1:07:49 PM

When we examine the A grid offenses, we find the relationships between plea rates and low-end sentences are all over the map. For example:

- Kalamazoo has a high rate of pleas for every offense type. It also has a high rate of low-end sentences for CSC and robbery but not for assault.
- Wayne has low rates of pleas for every offense type but high rates of low-end sentences.
- Washtenaw has a high rate of low-end sentences but its plea rates vary substantially by offense.
- Kent has higher plea rates for assault and robbery than its proportion of low-end sentences would suggest.

Table 20. Percentage of Conviction Method by Offense and County (Non-Habitual Sentences, 1999–2012)

County	Murder, Second-Degree			Assault with Intent to Murder			CSC, First-Degree			Armed Robbery		
	Plea	Court	Jury	Plea	Court	Jury	Plea	Court	Jury	Plea	Court	Jury
Berrien	76.2	0.0	23.8	57.1	0.0	42.9	74.4	0.0	25.6	94.2	0.0	5.8
Calhoun	60.0	0.0	40.0	60.0	0.0	40.0	67.6	1.4	31.1	83.8	0.0	16.2
Genesee	83.3	1.4	15.3	68.8	1.3	29.9	85.7	1.4	12.9	91.8	0.7	7.5
Ingham	85.7	0.0	11.9	52.2	0.0	47.8	62.9	1.6	35.5	79.7	1.3	19.0
Jackson	71.4	14.3	14.3	50.0	0.0	50.0	86.2	6.9	6.9	79.5	2.7	17.8
Kalamazoo	78.8	0.0	21.2	69.6	4.3	26.1	86.6	0.0	13.4	94.5	1.6	3.8
Kent	78.3	1.2	20.5	79.1	0.0	20.9	84.1	0.8	15.0	94.7	0.3	5.1
Macomb	54.3	0.0	45.7	63.0	0.0	37.0	68.5	0.0	31.5	88.6	0.8	10.7
Muskegon	72.2	0.0	27.8	63.6	9.1	27.3	82.7	3.7	13.6	93.6	0.0	6.4
Oakland	62.7	1.7	35.6	61.1	0.9	38.1	77.2	0.4	22.5	86.2	0.8	13.0
Saginaw	74.4	0.0	23.1	50.8	0.0	49.2	76.5	0.0	23.5	81.6	0.9	17.5
Washtenaw	76.5	0.0	23.5	73.1	0.0	26.9	75.0	0.0	25.0	91.1	1.1	7.8
Wayne	68.4	7.4	24.2	56.3	8.8	34.9	63.1	7.3	29.6	80.8	7.0	12.1
CMIS Data	Chi-Square $p < .001$			Chi-Square $p < .001$			Chi-Square $p < .001$			Chi-Square $p < .001$		

D. County-Specific Analyses

In addition to the full model and offense-specific models mentioned already, we ran separate analyses for each of the 13 counties, also found in Appendix B. This step allows us to further explore differences among counties that were noted earlier. Each county differed in the importance placed on individual guidelines factors and how the factors come together to explain sentence length. Breaking apart the counties is an important step in analysis, as Wayne County comprises a large portion of the sentences in our data, and therefore our overall model is more reflective of Wayne County which, as we will see, is unique in many ways.

When examining the variance model summaries by county of conviction, Wayne County stands out in several ways. First, the offense explained 34% of the variance in sentence length, the highest percentage of the 13 counties included. The remaining counties ranged from a low of 10% of variance explained in Calhoun County to 28% in Kalamazoo. In most of the counties the offense explained about 20–25% of the variance in sentence length.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Second, OV and PRV levels provided the lowest percent of variance explained in Wayne County at 11.9% and 17.6%, respectively.

- For OV level, the remaining counties ranged from the high of almost 39% of variance in Calhoun County to 12.2% in Ingham. For most of the counties OV level explained 25–30%.
- PRV Level explained almost 36% of the variance in Muskegon County down to 22% in Genesee County. For most of the counties PRV level explained 25–30%.

Further, while the full model and offense-based models show PRV level having greater explanatory power than OV level, this largely results from the influence of Wayne County. The PRV level explained more variance in only four counties (Ingham, Kent, Saginaw, and Wayne), while OV level also explained more variance in four counties (Berrien, Calhoun, Oakland, and Washtenaw). The remaining five counties had similar amounts of variance explained by OV and PRV levels. The differences among counties in the percentage of variance explained by habitual status displayed interesting results that will be examined in Chapter Fourteen.

Examining the amount of variance explained by specific guidelines factors reveals important differences among how counties utilize the sentencing guidelines overall. In the full model, guidelines factors explained 71% of the variance in sentence length; however, most counties have a greater amount of variance than that explained by these factors. Again, Wayne County is unique in that the guidelines factors explain the least amount of variance at 65.8%.

- At 66% and 70.7%, respectively, only Genesee and Saginaw Counties are also below the full model's level of explanation.
- Berrien County has the most variance in sentence length explained by the guidelines factors at almost 87%.
- Calhoun, Kalamazoo, Muskegon, Oakland, and Washtenaw are not far behind with at least 80% of variance explained.

A high amount of variance explained is evidence of sentence lengths being determined by guidelines factors. The three counties that have more than 30% of the difference in sentence length unaccounted for by guidelines factors suggest less adherence to the sentencing guidelines.

The inclusion of departures into the model reconciles a large amount of the differences among counties so that the percentage of variance explained ranges from 85% in Saginaw to 94.2% in Kalamazoo. Departures explain 23% of the variance in sentence lengths in Wayne County, which—as discussed previously—departs frequently. In contrast, in Calhoun County, departures explain 5.5% of the variance. In general, the counties that had relatively low levels of explanation from the guideline's factors have high percentages explained by departures, while the counties that had high levels of explanation from the guidelines factors had relatively low percentages of explanation from departures.

Turning to the impact of changes in PRV and OV levels on estimated sentence length, we again find variations by county. For example, Kalamazoo had some of the largest effects for increases in PRV level: the increase from Level A to B is a 57% increase, and at Level F it is 271% increase. Comparatively, in Genesee County, increases in PRV level had a more tempered effect on sentence length, with an increase of 23% for PRV Level B and 176% for PRV Level F.

County patterns in the effect on sentence length of OV level increases were less consistent. Moving from OV level II and then to OV level III**,** Wayne County had the smallest increases at 28% and 65%, respectively; Muskegon County had the largest increases, at 53% and 103%. For movement to OV level VI, the difference between the smallest and largest impact was 90 points. The sentence increased 189% in Saginaw compared to 279% in Washtenaw County.

The impact of departures on estimated sentence length varied among counties within a fairly narrow range. The impact of downward departures ranged from a 34% decrease in Calhoun County to

RECEIVED by MSC 4/12/2022 1:07:49 PM

more than 50% decreases in Jackson and Genesee Counties. Most of the counties saw an estimated decrease of 38–45%. The range is broader for upward departures. Ingham County had the smallest upward departure impact, with a sentence length increase of 63%, followed by Oakland and Washtenaw at 65% and 67%, respectively. In five counties, an upward departure was associated with more than a 90% increase. Interestingly, Wayne County, which we have previously highlighted for its extensive use of departures, fell in the middle for the impact of both downward and upward departures with 44% decreases and 70% increases, respectively.

Finally, the sentence increase for selecting a jury trial varied substantially by county. At the lower end, choosing a jury trial increased estimated sentence lengths by 11% in Ingham and Kent, 12% in Wayne, 15% in Kalamazoo, and 19% in Washtenaw. A jury trial added 24–26% to the sentence in Genesee, Macomb, Muskegon, and Oakland, and 30% in Calhoun. The most striking county disparities are the lack of any significant increase for a jury trial in Berrien and Jackson and, at the other extreme, a 38% increase in sentence length for having a jury trial in Saginaw.

E. Further Observations

While no data were available with regard to prosecution decision-making, a last point worth noting about counties is that each has one unique factor: the identity of the elected prosecutor. Because prosecutors are county officials who select the charges (including habitual offender enhancements), negotiate guilty pleas, agree or object to bench trials, recommend sentences, and decide whether to appeal downward departures, their influence on sentence lengths—both direct and indirect—is substantial, albeit difficult to measure. Some prosecution policies are idiosyncratic. They depend on the elected prosecutor's views about plea bargaining, sentencing, or particular offenses. However some practices may be a practical function of the county's size and thus may result in similar outcomes among like-sized jurisdictions. For instance, prosecutors in large counties with heavy caseloads, many personnel to manage, and a wide range of judges and defense attorneys to interact with may make more pragmatic, less individualized, and generally less publicly visible decisions than prosecutors in smaller counties.

County disparities in sentence length are often defended by judges and prosecutors on the basis of "local control" over "local crimes." The sentences are said to reflect differences in community values that local elected officials are bound to respect. Disparities are also presumed to reflect individualized decisions based on knowledge of local resources and circumstances.

The opposing view is that defendants are convicted in the name of the "People of the State of Michigan" of crimes defined by state statute. Their convictions and sentences are subject to appellate review by state-level courts, and their prison sentences are served in state facilities funded by all state taxpayers. Yet some Michigan prisoners are serving minimum sentences that are twice as long as those of other prisoners with similar prior records who committed similar crimes in a different county. This result is not only unfair to the individuals who drew the long straw but also means that some counties' sentencing decisions disproportionately affect the allocation of state resources.

In theory, the adoption of statewide sentencing guidelines indicates that the Legislature has resolved this dispute by coming down on the side of statewide consistency. However, in practice, the structure of the guidelines permits very substantial disparities among counties to continue unabated.

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER ELEVEN. SYSTEMIC FACTORS: SENTENCING JUDGE

Sentencing disparities do not only exist among counties. As Jesse Mack can attest, they also exist within counties among judges on the same bench.

> *On June 26, 2011, Jesse Mack, Christopher Hubbert, and Deangelo Reynolds went to a residence in Detroit to find Michael Miller, whom they believed stole a pistol from Mack's home the preceding week. Mack had reported the gun stolen without action by the police. Mack placed a shotgun in the trunk of the car because he believed Miller was at a "weed house." Mack drove. Miller was standing outside his residence. When Miller refused to talk to the occupants of the car and headed for his house, Reynolds fired the shotgun, which he had retrieved from the trunk, and Hubbert fired a pistol for a total of 14–15 shots aimed at the house. Miller was not hurt, but his mother, who was inside the house, sustained injuries to her forearm and calf.*
>
> *All three men were charged with assault with intent to commit murder, which carries a maximum sentence of life or any term of years, and possession of a firearm during the commission of a felony ("felony firearm"), which carries a mandatory sentence of two years to be served consecutively to the sentence for the underlying crime. Mack, who had not actually fired a weapon, was charged with both crimes as an aider and abettor. All pled no contest which, for most purposes, has the effect of a guilty plea.*
>
> *Mr. Mack, aged 23, had a high school diploma and various trade certifications, a well-paying job and no prior record, adult or juvenile. He lived with his longstanding girlfriend (whom he married while on bail) and their two children. He had no history of substance abuse or mental illness.*
>
> *On December 22, 2011, Judge Daniel Hathaway sentenced Mr. Mack to serve 7.5–20 years for the assault. The minimum was near the low end of the range recommended by Michigan's sentencing guidelines, which was 81–135 months (6.75–11.25 years). Mack also received two additional years for the felony firearm.*
>
> *The two co-defendants were not arrested immediately. They were sentenced nearly a year after Mr. Mack by a different judge after negotiating pleas with a different prosecutor.*
>
> *On November 14, 2012, Judge Gregory Dean Bill sentenced Mr. Reynolds. Because he had two prior felony convictions and was currently on parole, Mr. Reynolds's sentencing guidelines range was 225–375 months (18.75–31.25 years). He was sentenced to serve 5–15 years for the assault, plus two years for the felony firearm. His minimum of 60 months for the assault was 165 months—nearly 14 years—below the low end of the guidelines range. His Sentencing Information Report does not reflect any reasons for the departure.*
>
> *Six days later, Mr. Hubbert was also sentenced by Judge Bill. Like Mr. Mack, Mr. Hubbert had no prior record and his sentencing guidelines range was also 81–135 months. However, Mr. Hubbert was sentenced to 3–15 years for the assault and the additional two years for the felony firearm. His minimum of 36 months for the assault was 45 months—nearly four years—below the low end of the guidelines range. His Sentencing Information Report also does not reflect any reasons for the departure.*
>
> *Judge Hathaway, who had sentenced Mr. Mack, subsequently considered Mr. Mack's postconviction motion for relief from his sentence. Although Judge Hathaway thought the discrepancy in sentences between Mack and his co-defendants was "strange," he declined to question the decisions of other judges and concluded that Mr. Mack's sentence was proportional to the crime because it was within the range established by the sentencing guidelines.*

The discrepancies in this case are particularly stark because they occurred among co-defendants. But they are not unique. The data reveal that the median sentences of judges in the same county for the same type of offense imposed under the same legislative guidelines vary by three, six, 10, and even 12 years.

RECEIVED by MSC 4/12/2022 1:07:49 PM

The nearly 21,000 sentences in the CMIS dataset for 1999–2012 were imposed by hundreds of different judges. Many of those judges did not have a sufficient number of cases for analysis. Therefore, it is not possible to include "identity of the sentencing judge" as a variable in a regression analysis. We can, however, compare the sentences of those judges from a number of counties who did have a substantial volume of cases.

Wayne County, where Jesse Mack was sentenced, has the largest bench in the state. Table 21 shows us that if you were convicted of second-degree murder, you could expect to receive a minimum sentence of 16 years from Judge R, but a sentence of nearly 23 years if you were unfortunate enough to draw Judge B3. If you were convicted of assault with intent to murder, you were likely to receive a minimum sentence of 10 years if sentenced by Judge B1, but you risked receiving over 13 years more if you were sentenced by Judge B3. Similar disparities existed for the other offenses. Judges B2 and M had median minimum sentences of nine years for first-degree CSC while Judge B3.'s median sentence was four years longer. Judge A's 10-year median sentence for armed robbery was twice as long as the 5-year medians of Judges F and S.

Table 21. Median Non-Redundant Sentences Imposed by Select Wayne County Judges

Offense	No. of Judges	No. of Sentences	Low Median	Hi Median	Difference
Murder, second-degree	10*	567	16.0	22.75	6.75
Assault with intent to murder	7**	327	10.0	23.4	13.4
Criminal sexual conduct, first-degree	7***	310	9.0	13.0	4.0
Armed robbery	16****	3308	5.0	10.0	5.0

*More than 50 sentences each **At least 40 sentences each ***At least 39 sentences each ****At least 170 cases each

While other counties did not have the same volume of sentences, there were numerous examples of substantial disparities among judges who shared the same bench.

- In Oakland County CSC cases, Judge T typically imposed seven years while Judge G imposed 15 years. Among nine Oakland County judges who collectively imposed 855 armed robbery sentences, the median minimums showed a disparity of 8.5 years, ranging from 4.25 to 12.75 years.
- In Kent County, six judges had a difference of 1.5 years in their median sentences for CSC (from 10 to 11.5) and 2.8 years for their median robbery sentences (from 5.7 to 8.5).
- In Genesee County, while four judges imposed similar median sentences of either nine or 10 years for armed robbery, the same judges had very different medians for murder—18, 22.5, 24, and 30 years.
- In Saginaw County, three judges had median sentences for armed robbery that ranged from seven to 10 years, while the two Muskegon judges with enough cases for comparison had armed robbery medians of 10 and 10.3 years.

These findings, along with sample cases like Mr. Mack's, confirm what every criminal practitioner knows from experience: the sentence ultimately imposed depends, to a greater or lesser extent, on the luck of the draw. That image is not just a cliché; it is literally true. By court rule, circuit court cases are assigned to judges by lot.[1] This serves to even out judges' caseloads, avoid even the appearance of bias and prevent "judge-shopping" by the parties. Ironically, it also reinforces the randomness of sentence lengths. Whether the defendant is made to serve 10 years in prison or 15 may well come down to whether his case is assigned to Judge A in Courtroom 1 or Judge B next door in Courtroom 2.

RECEIVED by MSC 4/12/2022 1:07:49 PM

During interviews by a researcher, Michigan judges attributed differences in sentence lengths to varying judicial philosophies about the purpose of sentencing and the judge's role in the process. They also discussed the effect of local crime rates and local demographics, the need to work around the policies of different prosecutors, presentence report recommendations, the availability of alternative resources, and judges' attitudes toward such defendant characteristics as addiction, mental illness, family support, and employment.[2]

To the extent that it occurs in life-maximum cases, sentence bargaining can have some leveling effect. But experienced prosecutors and defense attorneys know the predilections of the local judges and what sentence agreements each judge is likely to accept. Thus the identity of the judge is itself a factor for consideration in negotiations. The parties also know how judges are likely to receive arguments for departures from guidelines recommendations, upward or downward.

The data show the final sentences—after plea negotiations and guidelines calculations have been taken into account by the identified judges. It is possible that without these constraints, the extent of disparity might have been even greater. It is undeniable that the sentencing guidelines do not prevent substantial disparities that are attributable not to what the defendant has done but to who the sentencing judge is.

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER TWELVE. DEFENDANT CHARACTERISTICS: SEX AND AGE

Disparity caused by systemic factors is difficult to address because it can result from policies that may be defensible or even seemingly inevitable. Plea bargaining, on which the operation of the criminal justice system has come to depend, is built on the idea that people ought to receive less punishment as a reward for waiving their right to a jury trial. Local officials justify county differences in sentencing on the rationale that local criminal justice outcomes should reflect local values and concerns. And differences in the sentences imposed by judges on similar defendants are said to be a byproduct of desirable judicial independence and experience.

Consideration of the personal characteristics of individual defendants presents a different situation. Except for the possibly mitigating factor of youth, demographic factors like sex, age, and race bear no relevance to culpability and have no justification in public policy or even the practicalities of administering public institutions. Given the over-representation of people of color among those who are incarcerated, the prospect of racial disparity, in particular, has long been a matter of concern.

Some defendant characteristics that might logically relate to sentence selection in particular cases include mental health, substance abuse history, employment status, and residential stability. We cannot examine their impact because this sort of background information was not part of the available data. However, prior research suggests that such factors are more likely to affect the In/Out decision for less serious offenses than the length of an inevitable prison sentence.

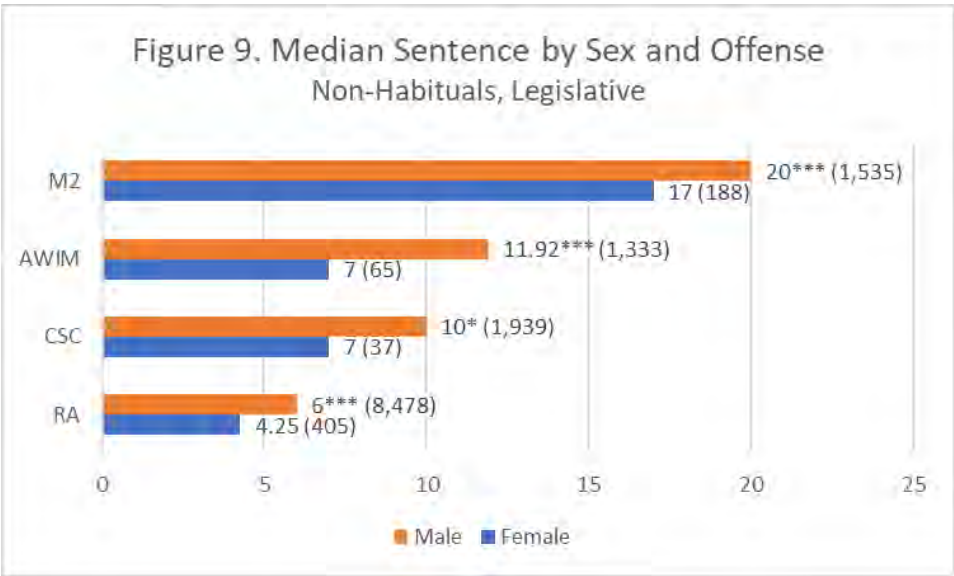
A. Sex

In 2019, women constituted 5.4% of the Michigan prisoner population. From 1999–2012, they accounted for 4.9% of the sentences imposed in life-maximum cases under the legislative guidelines. Their proportion of these sentences varied by offense group:

- Murder – 10.9%
- Assault – 4.6%
- CSC – 1.9%
- Robbery – 4.6%

Overall, at 69.7%, women are more likely than men, at 58.3%, to have only one sentence. This may reflect the fact that women are over-represented among murder sentences which are most likely to be single and under-represented among CSC sentences, which have a high proportion of multiples. Figure 9 shows substantial disparity by sex for non-habitual offender sentences in each offense group.[1]

RECEIVED by MSC 4/12/2022 1:07:49 PM



CMIS data, M2, AWIM, & RA*** $p < .001$, CSC* $p < .05$

Overall, 56.3% of the women's non-habitual sentences were at or below the low end of the applicable sentencing guidelines range, compared to 51.4% of the men's sentences.

Sex disparity may be largely the result of judicial perceptions that women are less dangerous than men. It may occur because women are often co-defendants with men who are viewed as the primary perpetrators. Or it may be because women simply have lower OV and/or PRV scores. Table 22 shows that women do, in fact, score substantially lower than men on prior record. With a few minor exceptions women are much more likely than men to score at Levels A–C; men score at Levels D–F at much higher rates than women.

Table 22. Percentage of Non-Habitual Sentences Among PRV Levels by Sex, 2003–2012

	Murder, Second-Degree		Assault w/Intent		CSC, First-Degree		Armed Robbery	
PRV Level	Female (N = 134)	Male (N = 1,084)	Female (N = 43)	Male (N = 841)	Female (N = 25)	Male (N = 1,200)	Female (N = 262)	Male (N = 5,394)
A	38.8	17.4	39.5	13.7	16.0	20.1	23.7	14.9
B	17.9	12.8	7.0	8.2	8.0	7.2	11.5	10.3
C	26.9	29.3	37.2	27.0	68.0	43.1	41.2	30.9
D	12.7	23.3	16.3	29.1	8.0	16.3	17.6	24.3
E	3.7	9.8	0.0	12.5	0.0	8.0	4.2	11.2
F	0.0	7.3	0.0	9.5	0.0	5.3	1.9	8.5
OMNI Data	$p < .001$		$p < .001$		Not Significant		$p < .001$	

The same level of sex disparity is not evident in the offense variable levels. No offense type showed statistically significant differences in offense severity between males and females. Another factor that may have helped nudge women's sentences below men's is that, for murder and assault, women's sentences were significantly more likely to result from guilty pleas.[2] Regression analyses also found little to no statistical difference in sentence length between the sexes. Exceptions to this were present in the offense-specific analysis for assault and the county analysis for Saginaw. In the assault regression, men were found to have an 11% increase in their sentence compared to women. In the county analysis, the results suggest that in Saginaw sentences were 13% longer for men.

RECEIVED by MSC 4/12/2022 1:07:49 PM

In sum, the sentences imposed on women were consistently shorter than those imposed on men by amounts ranging from nearly two to nearly five years. This discrepancy can be partially explained by women having lower PRV scores and a somewhat higher rate of guilty pleas.[3] However, given the size of the disparity, it seems likely that much of the difference is attributable to a combination of individual background factors not captured in this research and to more lenient attitudes toward female defendants.

B. Age

Table 23 examines the potential impact of age on sentences under the legislative guidelines. Overall, it shows no uniform relationship between the defendant's age and sentence length. There was not, for instance, a direct increase in median sentence as defendants got older. Although that was the general trend, there were numerous exceptions.

- The median sentence for murder was the same at virtually every age. The sole exception was juveniles younger than 17 whose median sentence was three years shorter.
- Those who were 17 and under had median sentences that, at below 11 years, were shorter than most in assault with intent to murder cases. Oddly the median dropped back to 10 years for those who were 40 or older.
- In the CSC and armed robbery groups, juveniles under 17 had longer median sentences than 17-year olds. The CSC group again shows a drop back to a 10-year median for defendants who were 40 and older.

In three offense groups, the largest number of sentences by far was imposed on people who were between 20 and 29. In the CSC group the most common age group was 30–39.

Table 23. Age at Offense and Median Minimum Sentence, Non-Habitual Sentences, 1999–2012

Age	Murder, Second-Degree		Assault with intent to Murder		CSC, First-Degree		Armed Robbery	
	Years	Number	Years	Number	Years	Number	Years	Number
<17	17.0	90	10.0	66	8.0	76	6.0	673
17	20.0	95	10.7	87	6.3	88	5.0	1,132
18–19	20.0	284	11.25	87	9.0	134	5.0	2,183
20–29	20.0	720	12.0	633	10.0	482	6.0	3,239
30–39	20.0	276	12.0	239	11.0	645	7.5	1,040
40+	20.0	258	10.0	165	10.0	551	8.0	616
Total	20.0	1,723	11.25	1,398	10.0	1,976	6.6	8,883
CMIS Data	$p < .05$		$p < .01$		$p < .001$		$p < .001$	

The age groups that show particularly interesting characteristics are under 17 and 17. Differences between the sentences for these groups and between the sentences for the entire OMNI dataset that were statistically significant at the $p < .001$ level appear below. Importantly, the law considered defendants younger than 17 to be juveniles who had to be waived into the jurisdiction of the circuit court in order to be tried and sentenced as adults.

- Those under 17 were less likely than those who were at 17 to score at Prior Record Levels A and B and more likely to score at PRV Levels C and D, indicating that the younger defendants had more severe juvenile histories. Not surprisingly, both groups of teenagers rarely accumulated enough of a criminal history to score at PRV Level E or F. The

RECEIVED by MSC 4/12/2022 1:07:49 PM

significantly different scoring in Wayne County of two particular prior record variables provides some insights.
    - Youth under 17 were almost twice as likely as those aged 17 to receive 10 points under PRV 3 for having a prior high severity juvenile adjudication.
    - On PRV 7, 33% of youth under 17 received 20 points for two or more concurrent or subsequent convictions compared to 23% of those aged 17.
- For the A grid crimes, those under 17 were less likely than those at 17 to score at Offense Levels 1 and 2. Their proportion at OV Levels 1 and 2 was also markedly below that of the group as a whole. Conversely, those under 17 were more likely to score at OV Levels 5 and 6 than either those who were at 17 or the group as a whole, suggesting that the younger defendants committed their offenses in a more serious manner.
- Both the youngest groups were more likely to plead guilty than defendants overall, which would militate toward shorter minimum sentences.
- Defendants under 17 had a markedly lower percentage of single sentences and a greater proportion of five or more sentences compared to the 17-year olds, again suggesting that their offenses were more serious.
- There was virtually no difference by age in the rate of departures. Apparently defendants' youthfulness did not move judges to be more lenient.

These findings suggest that the waived juveniles had more serious offenses and criminal histories than 17-year-olds convicted of the same crimes, which helps explain why the youngest defendants generally did not receive leniency based on their age—except perhaps when the crime was murder.

There were, however, significant differences among counties in the proportion of sentences attributable to waived juveniles.

- Saginaw accounted for 3.7% of all sentences but 8.3% of those under 17 as well as 5.3% of the 17-year olds.
- Less dramatically, Kent County accounted for 7.9% of all sentences but 10.3% of those under 17.
- Wayne County, on the other hand, accounted for 49.7% of all sentences and 52.8% of those imposed on 17-year-olds but only 40.7% of the under 17 group, which suggests that Wayne may be employing alternatives to treating juveniles as adults more often.

It has been suggested that guidelines commissions consider age as a mitigating factor at two extremes. Based on current brain science, younger defendants may be considered less culpable. Older defendants, on the other hand, are closer to "aging out" of crime. Even if they have substantial criminal histories, they are approaching the point where further offending is increasingly less likely. The length of their prior records, which may well consist of numerous non-violent offenses, could be balanced against the substantial cost of incarcerating older people.[4]

In general, the regression analyses did not have many significant findings related to age, except for confirming the lower sentences for those aged 17, compared to those both younger and older than 17. In the offense-specific models, if the defendant was 17 at the time of the offense, the sentence decreased by 10%. In the county-specific models, four counties (Berrien, Kalamazoo, Kent, and Macomb) showed that defendants who were 17 at the time of the offense received a decreased sentence. Additionally, in four counties (Calhoun, Macomb, Washtenaw, and Wayne), defendants older than 29 at the time of the offense saw a significant increase in sentence length.

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER THIRTEEN. DEFENDANT CHARACTERISTICS: RACE

Use of the term "non-white"

The Michigan Department of Corrections data on which this research is based does not distinguish the ethnic group "Hispanic" from the racial groups white and non-white. While race is a required field for probation officers who are inputting initial data, ethnicity is not. Hispanics, of course, have varied racial backgrounds depending on their country of origin.

The MDOC's Offender Tracking System (OTIS) shows that many people with common Hispanic surnames are identified as white while other people with identical names are identified as Hispanic. Because there is no reliable way to count Hispanics separately, they are simply included in whichever racial group they were placed at intake. Thus it is likely that many Hispanics are counted as white.

For the reader's information, Appendix Table C1 includes census data that shows the estimated percentage of Hispanics in each of the study counties.

There is no dispute that a disproportionate number of Michigan prisoners are non-white. In 2019, non-whites constituted 55% of the prisoner population although they accounted for only about 14% of the state's total population. For those convicted of life-maximum offenses, the numbers are even more skewed. Among all the sentences imposed under the statutory guidelines on people not sentenced as habitual offenders, non-whites accounted for 73.7% overall: 76% in the murder group, 78% in the assault group, 79% in the robbery groups and 44% in the CSC group. Of 2,300 sentences imposed from 1999-2012 on defendants who were 17 years or younger when they committed their crimes, 84% involved teenagers of color.

There are undoubtedly multiple reasons for these overwhelming differences. They may include higher arrest rates, higher actual crime rates, less favorable plea offers and less ability to post bail, investigate available defenses, or obtain adequate representation. In turn, these may result from economic, residential, educational, and health disparities rooted in historical and contemporary racism. However, the focus of the current research is limited to exploring the relationship of race to the sentencing guidelines and to the sentences ultimately imposed.

Many researchers have discovered just how complicated this exercise can be. Before examining the Michigan data, it is useful to consider Prof. Marie Gottschalk's summary of their efforts in *Caught*. Gottschalk quotes multiple researchers who have found little evidence of overt statewide racial disparity in sentencing but who suggest that differences in county-level policies may be masked by statewide data. Gottschalk goes on to express concern about the extent to which the research emphasis on racial and other forms of disparity has shifted focus from the policies that have caused harsher penalties for defendants of every race.[1] Referring specifically to the federal sentencing guidelines, she observes: "The most rigorous studies of the guidelines suggest that they have not had much success in reducing sentencing disparities between judges, but they have been wildly successful in raising the severity of sentences in most crime categories."[2]

Prior Michigan research also shows little or no racial disparity in sentence length when the analysis controls for prior record and offense severity. Recall that the Michigan Felony Sentencing Project found that 1977 sentences were significantly higher for non-white defendants only when the offenses were drugs, burglary, larceny and sex. In assessing sentences imposed under the judicial guidelines from 1992–1997, Ekpunobi found that race was significantly associated with sentence in only one of the four counties he studied. The National Center for State Courts found no significant difference by race in the average length of prison sentences.

RECEIVED by MSC 4/12/2022 1:07:49 PM

A. County Population

Before turning to the current research, it is useful to consider some basic figures about the non-white populations of the study counties compared to the proportion of sentences imposed on non-whites. These figures are summarized here. More details can be found in Appendix C.

In most of the study counties, the population is 18–26% non-white.

- Jackson is an exception at the low end with just 12.4% non-white.
- At the other end, Wayne County is unique, with a non-white population of 45.4%.

Five urban counties accounted for 88% of the non-white sentences, with nearly two-thirds coming from Wayne County alone. They were Wayne – 63.8%, Oakland – 8.3%, Kent – 6.4%, Macomb – 5.0%, and Genesee – 4.5%.

The proportion of sentences imposed on whites and non-whites varies by offense and county.

- Broadly speaking, the proportion of non-white sentences is two to three times larger than the proportion of non-whites in the county's population. So, for instance, 17.8% of Kent County's population is non-white, but 56% of its life-maximum sentences were imposed on non-whites.
- The proportion of white sentences is generally one-half or less of the proportion of whites in the county's population. Looking again at Kent County, 82% of its population is white but only 44% of its life-maximum sentences were imposed on whites.
- For Wayne County, the proportion of white sentences is just a quarter of the proportion of whites in the county's population.

When considered by offense type, there is notable variation.

- Only four counties have more white than non-white murder sentences (Calhoun, Ingham, Jackson and Kalamazoo).
- Only Muskegon has more white assault sentences.
- For every county except Wayne a higher proportion of CSC sentences are white than non-white.
- A higher proportion of every county's robbery sentences are non-white than white.

B. Age

The sentencing data reveals two fundamental points about race and age.

First, non-whites committed their offenses at a younger age than whites. The trend for every offense type is that, from age 17 on, the proportion of non-white sentences declines with each increase in age. For robbery, for instance, the non-white share of sentences is about 86% for those between 17 and 19 years old, 76% for those between 20 and 29, and then drops to 67% for both the 30–39 and 40+ groups.

- Among those sentenced for murder and assault, about 29% of the non-whites compared to 19% of the whites were younger than 20 at the time of their crimes.
- Among robbery defendants, 49% of the non-whites compared to 30% of the whites were younger than 20.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Second, the proportion of non-white sentences is much smaller for CSC than for other offense types at every age. It declines steadily from 61.4% for 17-year olds to 35.4% for those who are age 40 and over.

Table 24 breaks down for each offense type the percentage of sentences imposed on whites and non-whites by age group.

- For murder and assault, about 22% of all the sentences were imposed on non-white defendants who were 19 or younger at the time of the offense. Both racial groups show large spikes for these offenses at ages 20–29, and then decline. We cannot say how the circumstances under which these crimes were committed may relate to age.
- For robbery, which accounts for nearly two-thirds of all the life-maximum sentences, the figures shifted. The proportion of robbery sentences imposed on those who were 17 or younger was nearly twice the proportion imposed on these youth for murder and assault. More than 38% of all the robbery sentences were imposed on non-white defendants who were younger than 20 (compared to 6.3% imposed on white youth). The proportion again peaked at ages 20–29 but less sharply. The decline from there was even steeper. Armed robbery is clearly the crime that leads so many economically disadvantaged young men to prison.
- CSC shows a totally different pattern as to both age and race. The majority of sentences are imposed on white defendants. Relatively few of the offenses were committed by defendants younger than 20. The peak age group was 30–39. Defendants were 40 or older at the time of CSC offenses nearly four times as often as for murder and eight times as often as for robbery.

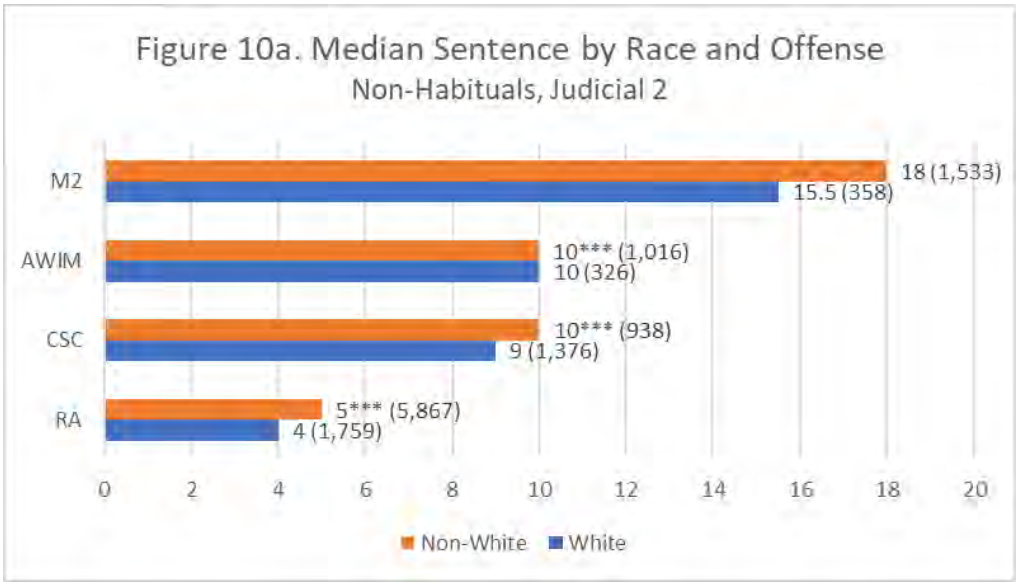
Table 24. Percentage of Sentences, by Offense Type, Age, and Race Non-Habitual Sentences, 1999–2012

	Murder, Second-Degree (N = 1,723)			Assault with Intent to Murder (N = 1,398)		
Age	Non-White	White	Total	Non-White	White	Total
17 & under	9.0	1.7	10.7	9.2	1.7	10.9
18–19	13.8	2.7	16.5	12.7	2.2	14.9
20–29	32.4	9.3	41.7	36.9	8.4	45.3
30–39	11.0	5.0	16.0	12.1	5.0	17.1
40+	10.2	4.8	15.0	7.4	4.4	11.8
Total	76.4	23.6	100.0	78.3	21.7	100.0
CMIS Data	p <. 001			p < .001		
	CSC, First-Degree (N = 1,976)			Armed Robbery (N = 8,883)		
Age	Non-White	White	Total	Non-White	White	Total
17 & under	4.6	3.7	8.3	17.7	2.6	20.3
18–19	4.0	2.8	6.8	20.9	3.7	24.6
20–29	12.4	12.0	24.4	27.7	8.8	36.5
30–39	13.5	19.1	32.6	7.9	3.8	11.7
40+	9.9	18.0	27.9	4.7	2.2	6.9
Total	44.4	55.6	100.0	78.9	21.1	100.0
CMIS data	p< .001			p < .001		

RECEIVED by MSC 4/12/2022 1:07:49 PM

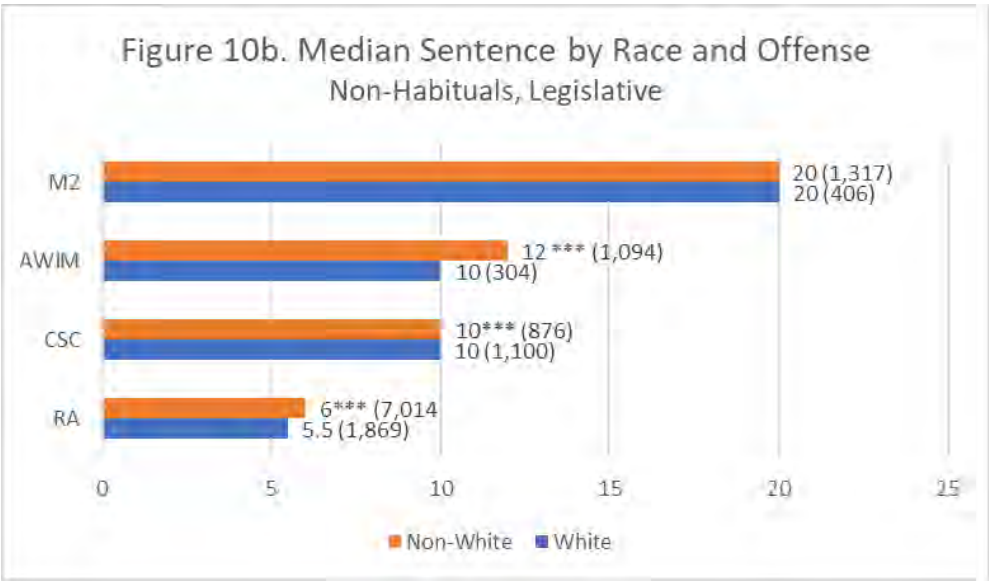
C. Median Sentence Length

The current research shows that when life-maximum offense groups are examined separately, statistically significant disparities in non-habitual offender sentences are apparent under both the judicial and legislative guidelines. Figure 10a shows that non-white defendants received significantly longer minimum sentences under the judicial guidelines for every offense except murder.[3]



CMIS Data, M2 not signif, AWIM, CSC, & RA *** p < .001
Means for white- M2 18.7, AWIM 10.5, CSC 11.5, RA 5.5
Means for non-white- M2 18.7, AWIM 13.0, CSC 14.75, RA 6.3

Figure 10b shows that this pattern of disparity continued under the legislative guidelines with the non-white sentences, except for murder, always being significantly longer.



CMIS Data, M2 not signif, AWIM, CSC, & RA *** p < .001
Means for white- M2 20.5, AWIM 12.0, CSC 10.9, RA 7.1
Means for non-white- M2 21.2, AWIM 14.1, CSC 12.6, RA 7.6

RECEIVED by MSC 4/12/2022 1:07:49 PM

The extent of the disparity changed between the judicial and legislative guidelines as median sentences increased for each racial group at different rates.

- White legislative sentences for murder "caught up" to non-white sentences because, while the median non-white sentence grew by two years, the median white sentence grew by 4.5 years.
- The disparity in CSC sentences declined because the median non-white sentence held at 10 years but the median white sentence grew by one year.
- Among robbery sentences the disparity decreased by six months because the white median sentence increased by 1.5 years while the non-white median increased by one year.

These changes suggest that the legislative guidelines may have reduced some disparity based on race. However, things went in the opposite direction for assault sentences. Where there had been no difference in the medians by race under the judicial guidelines, under the legislative guidelines the white median sentence held at 10 years but the non-white median grew by two years to 12.

As usual, aggregate figures mask substantial differences among counties.

- Murder sentences were significantly disparate in Ingham County ($p < .05$), where the median white sentence was 17 years but the median non-white sentence was 25 years.
- Assault sentences were significantly disparate in two counties, both at the $p < .05$ level.
  - In Oakland, the median sentence was 9.5 years for whites but 12 years for non-whites.
  - Similarly, in Wayne, the median was 9.25 years for whites but 11.25 years for non-whites.
- CSC sentences showed disparity in three counties, all at the $p < .001$ level.
  - The difference was most dramatic in Jackson, where the median sentence was 11.25 years for whites and 25 years for non-whites.
  - In Kalamazoo, the median for whites was just five years, while for non-whites it was 11.25 years.
  - In Wayne County, the median for whites was eight years, while for non-whites it was 10.5.
- Armed robbery had significant disparities in four counties.
  - In Muskegon ($p < .01$), white median sentences were five years while non-white sentences were 7.4 years.
  - In Oakland ($p < .01$), the difference was between whites at 4.25 years and non-whites at 5.5 years.
  - Wayne County ($p < .001$) had five years as the median for whites but six years as the median for non-whites.
  - In Genesee County ($p < .05$), it was non-whites with a median of seven years while the white median was nine years.

Three points stand out here. First, the low amount of racial disparity at the sentencing stage in most counties for most offenses is promising. Six counties (Berrien, Calhoun, Kent, Macomb, Saginaw, and Washtenaw) had no instances of statistically significant racial disparity; five others had disparity for one type of offense. Oakland's longer non-white sentences for both assault and robbery bear closer examination. The fact that Wayne County, which has by far the greatest number of sentences and the greatest number of non-white defendants, imposed significantly longer median sentences on non-whites is both unexpected and important to explore with more data than is available in the current research.

Second, of 10 instances of significant racial disparity, nine involved substantially longer sentences for non-whites. While not a surprise, this finding helps explain the greater number of non-whites incarcerated for life-maximum offenses.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Third, even where racial disparity is not apparent at the sentencing stage, more favorable treatment of whites at other stages may help explain the disproportionate number of non-white prisoners. A simple example would be if non-whites received less desirable plea agreements. Suppose, for instance, whites were more often allowed to plead down to lesser included offenses that did not carry life-maximum terms. Whites and non-whites who were convicted of similar offenses might show no evidence of disparity in their sentences but the greater proportion of non-whites sent to prison for life-maximum offenses would still be a product of racially disparate treatment. Or, in some counties, non-white juveniles may be waived into adult courts more often than white youth. The sentence lengths for whichever juveniles are waived may not differ by race but the proportion of non-white youth receiving adult prison terms would be the product of racial disparity. These are the sorts of hypothetical questions that merit further research.

It is important to note that, as used in the bivariate analysis, the variable "race" encompasses any factors that tend to be associated with race. The regression analysis holds other identified factors constant and focuses on the impact of race per se on sentence length. Counterintuitively, in the full model, if the sentence was for a defendant who was white, there was a 2% increase in sentence length. However, there were no significant findings for the impact of race alone on sentence length for any offense type or in any specific county.

Next we will examine identified factors that may indirectly cause racial disparities in sentence length because they impact non-white defendants differently than white defendants.[4]

D. Prior Record Level

Table 25 focuses on the PRV level. It shows that in three of the four offense groups, whites fell into Levels A, B, and C more often than non-whites, while non-whites fell into levels D, E, and F more often than whites. (It does not, of course, show that either group's prior convictions were for offenses similar to the current offense, only that non-whites had more of them.)

In the robbery group, this pattern was reversed. Non-whites showed less serious prior records than whites.

- Of the non-white sentences, 58.2% are at Levels A–C compared to 52.6% of the white sentences.
- Conversely, 7.2% of non-white sentences were at Level F compared to 11.6% of white sentences.
- Non-whites had roughly similar rates of sentences at each PRV level for robbery as they had for other offenses.

The shift in disparity occurred because white sentences were scored at PRV Level A about half as often for robbery as they were for the other offense types. White robbery sentences were scored much more often at PRV Levels E and F.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 25. Percentage of Non-Habitual Sentences Among PRV Levels by Race and Offense, 2003–2012								
	Murder, Second-Degree		Assault w/Intent		CSC, First-Degree		Armed Robbery	
PRV Level	White (272)	Non-white (946)	White (205)	Non-White (679)	White (703)	Non-White (522)	White (1,197)	Non-White (4,462)
A	27.2	17.7	22.4	12.7	24.0	14.6	12.6	16.0
B	14.7	13.0	10.2	7.5	7.7	6.7	12.4	9.8
C	32.7	28.0	31.7	26.2	50.2	34.7	27.6	32.4
D	15.8	24.0	22.4	30.3	12.1	21.6	24.6	23.8
E	7.0	9.7	6.8	13.4	4.1	12.8	11.2	10.8
F	2.6	7.6	6.3	9.9	1.8	9.6	11.6	7.2
OMNI Data	p <.001		p <.001		p <.001		p <.001	

When we examine racial differences in PRV Levels by county we find that non-whites had significantly higher scores than whites in Wayne ($p < .001$), Oakland ($p < .05$), and Muskegon ($p<.001$).

- Among Wayne County sentences, 34.1% of whites but only 25.8% of non-whites scored at PRV Levels A and B. However, 45.4% of white sentences scored at combined levels C and D compared to 52.5% of non-white sentences. This finding helps explain the higher median sentences for non-whites in Wayne. Notably, there was little difference between racial groups in the proportions scored at Levels E and F.
- Among Oakland County sentences, the important differences occurred at PRV Levels C and D. The proportion of non-white scores was 8.4 points lower than white scores at Level C, and 8.3 points higher at Level D. This finding may help explain Oakland's higher non-white sentences for assault.
- Muskegon has dramatic differences by race, with 88.7% of whites but only 59.8% of non-whites scoring at PRV Levels A–C. It must be noted, however, that the overall number of Muskegon sentences at these levels are relatively small (white = 58, non-white = 65), so that percentage differences appear larger.

To try to understand how non-white prior record scores come to be higher, we examined the scoring of each individual prior record variable. Table 26 provides a number of observations.

- For both high and low severity felonies, non-whites had significantly fewer sentences with zero points for every offense type except robbery. The differences for robbery were much less extreme but nonetheless they showed a larger proportion of zero points for non-whites. This finding may reflect the relatively younger age of non-white robbery defendants who would have less time to accumulate adult felony convictions.
- High severity juvenile adjudications showed very little difference by race; the vast majority of all scores showed zero points. Low severity juvenile adjudications showed non-white sentences with a significantly lower share of zero point scores except for robbery offenses. Even for robbery, non-whites had a lower share but the difference from whites was small and not statistically significant.
- A much lower proportion of sentences for both racial groups had zero points for misdemeanors. There was a significant difference for every offense type except CSC, but the direction of the results was not consistent. Non-white sentences showed a larger share of zero

RECEIVED by MSC 4/12/2022 1:07:49 PM

points for assault and robbery; white sentences showed relatively more zero point scores for murder.

- There was a highly significant difference between racial groups in scores for their relationship to the criminal justice system. Once again, robbery went in a different direction than the other offense types with a higher percentage of non-white sentences being scored at zero points.
- PRV 7 showed no significant differences by race for any offense type. With a few exceptions, PRV 7 generally showed a much lower proportion of zero point scores than the other PRVs for both whites and non-whites.

Table 26. White vs. Non-White Percentage of Sentences with Zero Points for Each Prior Record Variable Non-Habitual Sentences, 2003–2012

		Murder, Second		Assault		CSC, First		Robbery	
		White	N-W	White	N-W	White	N-W	White	N-W
1	High Severity Felonies	89.7***	79.4	85.9*	75.3	92.3***	75.3	78.9*	82.0
2	Low Severity Felonies	77.9***	54.8	74.1***	51.3	84.8***	66.3	60.9***	65.2
3	High Severity Juvenile	96.3 ns	92.4	95.1*	91.0	95.6 ns	93.7	91.7 ns	89.1
4	Low Severity Juvenile	87.9**	77.9	89.8***	75.3	96.3***	89.1	79.7 ns	76.0
5	Misdemeanors	55.9***	54.3	49.3**	55.7	66.7 ns	59.0	36.7***	58.2
6	Relationship to Criminal Justice System	74.3***	61.5	66.8***	59.1	88.3***	72.4	59.1***	61.0
7	Subsequent or Concurrent Felonies	67.6 ns	64.3	53.2 ns	45.1	45.4 ns	41.0	53.8 ns	52.4

OMNI data, *p < .05, **p < .01, ***p < .001, ns = not significant

Overall, it appears that higher non-white prior record levels are primarily attributable to PRVs 1, 2, 4, and 6. Although non-white sentences tended to have lower scores on some PRVs among robbery sentences, the differences in those scores, although statistically significant, were much less extreme than the differences between non-whites and whites for other offense types.

Racial disproportionality in sentencing resulting from higher non-white criminal history scores is not unique to Michigan. The Robina Institute has suggested that policymakers weigh the risk-predictive value of each prior record variable against added racial disparate impact. If a variable has little or no ability to predict future offenses—particularly assaultive offenses—but it contributes substantially to racial disparity, the variable may be eliminated or its weight reduced. Sentencing commissions may also wish to consider dropping or reducing the weight accorded to particular kinds of prior offenses, such as drug crimes, that play a relatively greater role in raising the criminal history scores of non-white as opposed to white defendants.[5]

E. Offense Severity Level

Table 27, which focuses on OV Level, shows more variation among the offense groups.

- There was virtually no difference in OV level by race among murder sentences.
- The assault group presented the largest and most telling difference. The disparity at Level VI, where the percentage of non-whites was 12.3 points higher than the percentage of whites, helps explain the two-year racial difference in median sentence lengths for assault.
- The results in the CSC group were somewhat uneven. At Levels I and II, the proportion of non-white sentences was 17.6%, while the proportion of white sentences was 13.5%. Conversely, 42.7% of the white sentences fell into Levels IV and V compared to 33.6% of the

RECEIVED by MSC 4/12/2022 1:07:49 PM

non-white sentences. This finding suggests a tendency for non-whites to have lower OV levels than whites. However, non-whites' percentage of sentences at OV Level VI was nearly five points higher than whites'.

- Among the armed robbery sentences, whites showed a clear tendency to have lower OV scores. Nearly 58% of the white sentences scored at OV Levels I and II compared to 45.5% of the non-white sentences. Notably, most of the difference occurred at Level I.

Table 27. Percentage of Offense Variable Level by Race and Offense Non-Habitual Sentences, 2003–2012

	Murder, Second		Assault w/ Intent		CSC, First		Armed Robbery	
OV Level	White (272)	Non-white (946)	White (205)	Non-White (679)	White (703)	Non-White (522)	White (1,197)	Non-White (4,462)
I	6.3	6.9	2.9	1.6	1.8	3.4	20.0	9.7
II	41.5	39.3	5.9	2.1	11.7	14.2	37.8	35.8
III	52.2	53.8	13.7	8.1	16.6	16.9	25.0	28.6
IV	---	---	23.4	20.0	24.8	19.0	9.4	13.8
V	---	---	19.0	20.8	17.9	14.6	3.5	5.4
VI	---	---	35.1	47.4	27.2	32.0	4.4	6.7
OMNI Data	Not Significant		$p < .001$		$p < .05$		$p < .001$	

F. Conviction Method

Another key factor that displays differences by race is conviction method. Table 28 shows that in every offense group, whites have a significantly higher proportion of guilty pleas than non-whites. Conversely, non-whites had higher rates of both jury and bench trials. We know that sentences imposed after both types of trial are longer than those that result from a guilty plea. Thus the difference in conviction method is directly associated with longer sentences for non-whites.

Table 28. Percentage of Conviction Method of Non-Habitual Sentences, 1999–2012

Crime	White			Non-white		
	Plea	Court	Jury	Plea	Court	Jury
Murder, 2nd	75.6	2.2	22.2	69.2	6.3	24.5
Assault	68.4	2.0	29.6	56.7	6.8	36.6
CSC, 1st	78.3	1.1	20.6	64.0	6.3	29.7
Robbery	91.7	1.7	6.6	82.8	4.9	12.3
CMIS Data, All offense types significant at $p < .001$						

Why non-whites have fewer guilty pleas is unclear. It could be that non-whites in general are offered less attractive plea bargains or are more resistant to pleading or have less effective defense counsel. It may also simply result from the fact that the majority of non-whites are convicted in Wayne County which, as we have seen, tends to have a lower rate of guilty pleas and a higher rate of both jury and bench trials than other counties.

RECEIVED by MSC 4/12/2022 1:07:49 PM

G. Place on the Range

Higher prior record and offense variable scores can explain why non-white sentences would be chosen from higher ranges on the applicable grid. But once the appropriate range is identified and the trial judge must select a sentence that is within, below or above that range, do further racial disparities appear?

Our examination of the data produced intriguing results.

- Place on the range showed no significant differences by race for most counties.
- Statistically significant differences existed for four: Genesee ($p < .01$), Kent ($p < .05$), Oakland ($p < .001$), and Saginaw ($p < .001$).
- In each of these four counties, non-white sentences are more likely to be both below and at the minimum-minimum than white sentences.
- In three of these counties, white sentences are more likely to be at or above the maximum-minimum. In Saginaw, white sentences are more likely to be above the max-min.

Table 29 shows the proportions by race of low-end sentences [downward departures + sentences at the min-min] and of high-end sentences [sentences at the max-min plus upward departures]. The differences are substantial.

Table 29. Race and Place on Range: Four Counties
Non-Habitual Sentences, 2003–2012

County	Low End		High End	
	White	Non-white	White	Non-white
Genesee	28.7% [20.5 + 8.2]	43.1% [31.5 + 11.6]	32.9% [24.0+8.9]	21.5% [18.6+2.9]
Kent	42.4% [33.0 + 9.1]	47.4% [37.9 +9.5]	10.6% [6.8+3.8]	5.2% [2.5+2.7]
Oakland	48.1% [13.0 +35.1]	55.8% [18.5+ 37.3]	11.1% [7.4+3.7]	5.8% [4.2+1.6]
Saginaw	26.6% [19.0 + 7.6]	44.8% [20.9 +23.9]	25.3% [7.6+17.7]	15.9% [14.1+1.8]

The various possible explanations for this phenomenon are not necessarily the same for each county. For example, judges may have felt that the guidelines scores produced especially harsh sentences for non-whites. They may have wanted to compensate for what they saw as less favorable plea bargains for non-whites. Or, on the contrary, more low-end sentences for non-whites may have been the product of bargains for certain sentence recommendations. Whatever the cause, in these four counties, non-white defendants had a better chance than whites of receiving a sentence at or below the minimum of the applicable range.

A similar situation exists in Macomb County. The distribution of sentences across the entire range is not statistically significant in Macomb. However, there is a notable difference right at the minimum-minimum, where 33.9% of non-white sentences but only 24.4% of white sentences fell. That is, while there is not a lot of difference by race in general, non-whites were substantially more likely than whites to receive a sentence right at the bottom of the applicable range.

Wayne County presents an analogous but opposite picture. The distribution of sentences across the entire range is also not statistically significant in Wayne. However, there is a notable difference in the rate of downward departures. While sentences below the min-min are common for everyone in Wayne County, they are more common for whites, at 51.5%, than for non-whites, at 46.9%. This lower rate of

RECEIVED by MSC 4/12/2022 1:07:49 PM

downward departures may help explain why non-white sentences in Wayne County are longer than white sentences.

H. Conclusion

The discussion of defendant characteristics and of race in particular returns us to where we started: these issues are enormously complex. Multiple factors are intertwined, and separating cause from effect is difficult to do with any certainty. Undesirable consequences may have been unintended or accepted as the cost of strategic choices. The possibility of untested influences that relate to multiple factors is always present. If significant disparities in sentence length by sex, age, and race are attributable, at least in part, to the operation of the guidelines, difficult policy choices must be recognized.

The findings regarding Wayne County illustrate these issues. Except in the case of murder, Wayne's median sentences are often longer than those of other large counties. Wayne County accounts for nearly two-thirds of all the non-white sentences, whites' sentences are shorter than those of non-whites, and non-white sentences reflect relatively fewer downward departures. Does all this mean that Wayne County sentences are the result of racial discrimination?

We also know that Wayne County sentences show significantly higher prior record levels for non-whites than whites, that non-whites tended to score at higher offense levels, and that non-whites pleaded guilty less often than whites for reasons that are unclear. Could those racial differences be enough to explain the county's sentencing disparities? Could non-white higher prior record scores in themselves be evidence of racial disparity in adult arrest, charging, and bargaining decisions and/or the treatment of juveniles? Could they result from how the prior record variables are designed or scored? Or might Wayne County's higher PRV scores for non-habitual offender sentences reflect low rates of habitualization so that more people with long prior records appear in the non-habitual sentence data than is true for some other counties?

These questions do not even begin to address differences among counties that this research cannot assess. But they do identify potentially fruitful areas for future research. The data suggest other intriguing questions as well. Why are median sentences for CSC longer for non-whites even though they have a much lower proportion of sentences for CSC than for any other offense? Do the CSC offenses committed by non-whites and whites tend to differ somehow? Are there differences in sentence length related to the race of the victim? Why are prior record levels less serious for armed robbery offenses? Is it just a function of the crimes being committed by younger defendants? Why do non-whites have higher scores for offense severity? Are all the differences in scoring by race objectively supported, or should they be examined for implicit bias?

The bottom line is that identifying the existence of disparity is easier than explaining its causes. Race and county are both shorthand for a constellation of interacting characteristics that include the operation of the guidelines themselves. Ultimately, we will discuss ways to reduce sentencing disparity regardless of its causes.

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER FOURTEEN. HABITUAL OFFENDER ENHANCEMENTS: WHO GETS CHOSEN?

Chapters One and Five discuss the purpose of the habitual offender statutes, how they were integrated into the legislative sentencing guidelines, and what the consequences have been. In this chapter, we will examine which of the defendants who qualify for a habitual offender enhancement actually receive one.

Because the decision to habitualize is made by the prosecutor, it is not surprising that practices vary by county. In its analysis of 2012 data, the Council of State Governments (CSG) found that 42.4% of defendants whose prior convictions made them eligible for enhancements were actually sentenced as habitual offenders—an increase from 2008, when the proportion was 38.5%.[1] CSG found enormous differences in the use of habitual enhancements by county. It ranged from 10% of those eligible in Washtenaw County to 89% of those eligible in Oakland County. Wayne, Kent, and Macomb Counties were all below 25%. The other county with an enhancement rate over 85% was Saginaw.[2]

The Criminal Justice Policy Commission (CJPC) also undertook an analysis of habitual offender sentencing, using data from 2012–2017. Although the CJPC was disbanded before the analysis was completed, its preliminary findings included considerable geographical differences. The counties that applied the enhancement most frequently were Muskegon, Oakland, and Saginaw.[3]

Because of the prominence of this finding, which we will confirm shortly, it is important to bear in mind that these research results may or may not reflect any given county's current practices. Because the rate of habitualization reflects the policies of the elected prosecutor, changes in either direction can occur when a new prosecutor is elected.

A. How Many Sentences are Eligible for the Habitual Offender Enhancement and How Many Receive it?

As discussed In Chapter Five, under the legislative guidelines 18% of sentences included in our research received a habitual offender enhancement. To understand how many sentences were eligible for enhancement, we used the OMNI data to create a separate dataset that only included sentences that were habitualized or had points for PRVs 1 and 2. Scores for PRVs 1 and 2, which count prior high and low severity adult felony convictions, are utilized as a proxy for eligibility. This is a conservative estimate of the number of sentences eligible for enhancement, as scoring PRVs is subject to the 10-year gap rule, while there is no time restriction on the use of prior convictions for habitual enhancements.[4] For example, included in this subset there are 26 habitual sentences that scored at PRV level zero, meaning none of their prior convictions could be counted in the scoring of the PRVs, but could still be used in the decision to habitualize.

Table 30 displays by offense type the percentage of sentences that were enhanced out of the total sentences eligible:

- Murder is habitualized the least often, with less than 25% of eligible sentences receiving the enhancement.
- Assault sentences receive the habitual enhancement for about 32% of those eligible.
- CSC sentences have the highest rate of habitualization, at almost 45% of eligible sentences receiving the enhancement.
- Robbery sentences receive the habitual enhancement at a rate similar to assault sentences, at 34% of those eligible. Due to the large number of robbery sentences, in sheer numbers most of the habitual sentences in our data are for robberies.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 30. Percentage of Habitual-Eligible Sentences Enhanced, by Offense, 2003–2012		
	Not Enhanced	Enhanced
M2	75.3 (551)	24.7 (181)
AWIM	68.3 (441)	31.7 (205)
CSC	55.3 (352)	44.7 (285)
RA	65.8 (2,369)	34.2 (1,229)
Total	66.2 (3,713)	33.8 (1,900)
OMNI Data, $p < .001$		

The preliminary findings of the CJPC were similar, finding that 24.1% of second-degree murder convictions, and 31% of convictions on the A-grid received the enhancement.[5]

Like the researchers from CSG and the CJPC, we examined enhancement rates by county. Our results were similar—albeit more extreme—with only 13% of Wayne County sentences and 91% of Saginaw's being enhanced. Table 31 displays the percentage of eligible sentences that received enhancement by county.

Table 31. Percentage of Habitual-Eligible Sentences Enhanced, by County, 2003–2012		
County	Not Enhanced	Enhanced
Berrien	57.1 (48)	42.9 (36)
Calhoun	31.1 (32)	68.9 (71)
Genesee	48.9 (129)	51.1 (135)
Ingham	55.7 (88)	44.3 (70)
Jackson	48.3 (28)	51.7 (30)
Kalamazoo	55.3 (83)	44.7 (67)
Kent	75.4 (337)	24.6 (110)
Macomb	81.3 (278)	18.7 (64)
Muskegon	8.3 (12)	91.7 (133)
Oakland	16.6 (118)	83.4 (591)
Saginaw	8.3 (17)	91.7 (188)
Washtenaw	70.5 (98)	29.5 (41)
Wayne	87.0 (2,445)	13.0 (364)
OMNI Data, $p < .001$		

The counties can be roughly divided into three groups based on their rate of enhancement:

- Low (13%–30%): Kent, Macomb, Washtenaw, and Wayne Counties
- Moderate (42%–52%): Berrien, Genesee, Ingham, Jackson, and Kalamazoo Counties.
- High (more than two-thirds of eligible sentences): Calhoun, Muskegon, Saginaw, and Oakland Counties.

The rate of habitualization does not seem to be associated with location or size. Wayne, Oakland, and Macomb, the three large counties in southeast Michigan represent both extremes. Muskegon, Saginaw, and Oakland, which have the highest rates of habitualization, are different sized counties in different parts of the state.

RECEIVED by MSC 4/12/2022 1:07:49 PM

B. How are Sentences Selected for the Habitual Enhancement?

1. Guidelines Factors

While prior record scoring and habitual offender enhancement inevitably overlap to the point of double counting, we hypothesized that the decision to enhance may be influenced by just how extensive the defendant's criminal history is. That is, the higher the prior record level, the more likely the sentence will be enhanced. Table 32 displays the percent at each PRV level for habitual-eligible sentences that were and were not enhanced.

Table 32. Percentage of Habitual-Eligible Sentences Enhanced, by PRV Level, 2003–2012

PRV Level	Not Enhanced	Enhanced
A (0)	0	1.4
B (1–9)	4.1	1.1
C (10–24)	16.3	6.1
D (25–49)	38.2	25.2
E (50–74)	23.2	33.5
F (75 +)	18.2	32.8
OMNI Data, $p < .001$		

Overall, we can see in the above table a much higher proportion of enhanced sentences are at higher PRV levels. In the sentences not enhanced, almost 60% scored at level D or below. The opposite is true for the sentences that did receive the habitual enhancement: more than 66% scored in the highest two PRV levels. The same analysis was conducted for each county and found that:

- PRV level was not significant in four counties (Berrien, Jackson, Muskegon, and Saginaw). These counties were identified as moderate-to-frequent users of the habitual enhancement, including the two counties that utilize the habitual enhancement the most frequently.
- Differences in PRV level for those selected for habitual enhancement was significant in the remaining eight counties.

Similarly, it could be anticipated that sentences with the highest OV levels would be selected for habitualization. However, analysis did not find this to be true. In examining the distribution of habitual sentences by OV level for the counties together, there was no significant difference for murder or CSC offenses. The remaining two offense types were significant, however there was no discernable pattern. There were also no discernable patterns when examining the OV level by county.

Beyond the guidelines factors of offense, PRV, and OV level, systemic factors and defendant characteristics could help explain the difference in which sentences are enhanced and which are not.

2. Conviction Method

Besides county, the systemic factor that may relate to enhancement is conviction method. Table 33 shows a significant difference in conviction method between those who were eligible but did not receive the enhancement and those that did.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 33. Percentage of Habitual-Eligible Sentences Enhanced, by Conviction Method, 2003–2012		
Method	Not Enhanced	Enhanced
Plea	79.4	51.6
Court	4.6	6.5
Jury	16	41.9
OMNI Data, $p < .001$		

Almost 80% of the sentences that were not enhanced were settled by plea—compared to just over half of the habitualized sentences. As indicated earlier, the threat of habitual enhancement can be used in plea negotiation. The evidence of greater plea rates for sentences that were not habitualized supports this possibility. However, there were obviously still many people who pleaded and received enhanced sentences, nonetheless. We do not have data that would indicate whether negotiations may have resulted in enhancement at a lower level such as habitual second instead of habitual fourth.

Analysis of conviction method by county, found that:

- There was significant difference in the method of conviction for sentences that did and did not receive the habitual enhancement in 10 counties.
- The three counties that did not have a significant result for conviction method are the counties that utilize the enhancement the most often—Muskegon, Oakland, and Saginaw. It appears that in these counties the policy was to enhance, at least to some degree, in most eligible cases regardless of the defendant's willingness to plead.

3. Defendant Characteristics

In previous chapters we examined the relationship of sex, age, and race to median sentence length. Examining the relationship of these defendant characteristics to habitual offender status presents different issues. There were too few women with enhanced sentences to make the analysis meaningful.[6] The age of the defendant at the time of the offense is significantly related to the application of the habitual enhancement, but that would be expected since those who were younger have had less time to acquire adult felony convictions.

That brings us to race. The overall analysis found that 39.4% of sentences for white defendants received the habitual enhancement, compared to 32.1% of sentences for non-white defendants. That is, relatively more sentences committed by white defendants were enhanced. Moreover, in seven of the 11 counties where no statistically significant relationship was found more white sentences received the enhancement. In the counties where there was a significant relationship between race and habitual enhancement, the opposite was true. In Berrien County, 54.5% of eligible sentences committed by non-white defendants received the sentencing enhancement, compared to 20.7% of sentences with white defendants. The difference in Genesee County was less extreme, with 55.7% of sentences for non-white defendants, and 40.5% of sentences for white defendants receiving the sentencing enhancement. In the preliminary findings of the CJPC, they found more non-white defendants were eligible for the habitual enhancement but received the enhancement at a similar rate as white defendants.

C. How Much Longer are Habitual Sentences?

Table 34 shows the median minimum sentence length in each county for habitual-eligible sentences by their habitual status. The increase in median sentence length depending on habitual status was significant in every county except for Jackson.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 34. Median Sentence Length in Months for Habitual-Eligible Sentences by County and Habitual Level, 2003–2012					
	Not Enhanced	2nd Habitual	3rd Habitual	4th Habitual	Increase between Not and 4th Habitual
High (More than two-thirds of eligible sentences)					
Calhoun***	180 (32)	240 (27)	300 (21)	420 (23)	133%
Muskegon***	134 (12)	144 (51)	228 (29)	288 (53)	115%
Oakland***	120 (118)	135 (183)	150 (147)	180 (261)	50%
Saginaw*	120 (17)	206 (88)	228 (41)	225 (59)	87.5%
Moderate (42%–52% of eligible sentences)					
Berrien***	190 (48)	279 (13)	240 (9)	405 (14)	113%
Genesee***	150 (129)	192 (40)	262.5 (32)	240 (63)	60%
Ingham**	141 (88)	135 (29)	200 (12)	240 (29)	70%
Jackson ns	180 (28)	250 (10)	240 (11)	300 (9)	60%
Kalamazoo***	120 (83)	216 (33)	318.5 (12)	292.5 (22)	144%
Low (13%–30% of eligible sentences)					
Kent***	120 (337)	240 (42)	177 (26)	270 (42)	125%
Macomb***	108 (279)	126 (31)	171 (12)	180 (20)	67%
Washtenaw***	126 (98)	288 (10)	348.5 (18)	180(13)	43%
Wayne***	120 (2,446)	156 (109)	240 (85)	240 (169)	100%
Total***	120 (3,715)	169.5 (666)	216 (455)	224 (777)	87%
OMNI Data, *** $p < .001$, ** $p < .01$, and *$p < .05$, ns = not significant					

The extent to which sentences got longer as habitual status increased varied widely. For example:

- In moving from no enhancement to second habitual status, the sentence increase was just 10 months for Muskegon, 15 months for Oakland, and 18 months for Macomb. By comparison, the increases were 86 and 89 months in Saginaw and Berrien, respectively, and a whopping 10 years for Kent.
- In moving from second to third habitual status, Oakland again showed a modest increase of 15 months. On the other hand, Wayne's median increased by 84 months, Kalamazoo's median increased by 102 months, while Kent's actually declined by 63 months.
- In moving from third to fourth habitual status, Calhoun's median increased by 10 years and Berrien's increased by almost 14 years, while Macomb's increased by a modest 9 months, Wayne's stayed the same and Genesee's declined.

The ultimate difference between no enhancement and being enhanced as a fourth offender varied tremendously. It appears unrelated to how often the county used the enhancement.

- In five counties the fourth habitual sentence was at least double the median sentence for those not habitualized.
- In six counties, the increase between sentences that were not habitualized and fourth habituals ranged from 60% to 87.5%.
- Kent and Wayne Counties applied the habitual enhancement to fewer than 25% of eligible sentences; but when they did, the increase in sentence length was substantial.
- Oakland County was unique in that it utilized the habitual offender enhancement often but had the smallest increases in sentence length. The change in sentence length between not habitualized and fourth habituals was 50% or 60 months (five years).

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 34 reveals several other points.

- There was no consistent pattern of sequential increases from one habitual level to the next. Sometimes a county's median sentence for the third offender was higher than its median for the fourth offender. The median for second offenders might be lower than the median of sentences that were not enhanced or higher than those of third or even fourth offenders.
- The fact that a county's non-enhanced sentences were relatively low or relatively high did not mean that its enhanced sentences were similarly low or high.
- The status of third habitual offender was generally used substantially less than either habitual second or habitual fourth, suggesting that many habitual fourth enhancements were being plead down to habitual second or third.

Earlier we suggested that differences in non-habitual sentence lengths could be related to habitual sentencing enhancement practices. A county that uses enhancements routinely will move many sentences with higher PRV scores into its habitual group; while in a county that enhances less frequently, the sentences with higher PRV scores will be in the non-habitual group, pushing up that group's median sentence. To test this hypothesis, we compared the armed robbery sentences of Oakland County, which enhanced sentences routinely, to those of Wayne County, where enhancements were infrequent. As we saw in Table 17, Oakland's median non-habitual sentence was five years, and Wayne's was six.

First, we divided the sentences for each county into three groups: those not eligible for enhancement, those that were eligible but did not receive the enhancement, and those that were enhanced. We then focused on the first two groups.

For those not eligible for enhancement, we multiplied the number of sentences by the median sentence to get the total number of months. For the eligible but not enhanced sentences, we also multiplied the number of sentences by the median sentence to get a total number of months. We then added together the number of months for each group and added together the number of sentences for each group and then divided months by sentences to get a new median for the combined groups. When we applied this formula to Wayne and Oakland Counties, it looked like this:

In Wayne County:
For the sentences not eligible for the enhancement,
51 months x 1,539 = 78,489
(median) (number of sentences)
For the eligible sentences,
84 months x 1,488= 124,992
(median) (number of sentences)

(78,489 +124,992) ÷ (1,539 +1,488)= 203,481/3,027= **67 months**

In Oakland County:
For the sentences not eligible for the enhancement,
54 months x 484 = 26,136
(median) (number of sentences)
For the eligible sentences,
108 months x 90= 9,720
(median) (number of sentences)

(26,136 + 9,720) ÷ (484+90)=35,856/547= **62.5 months**

Thus, it appears that the hypothesis is accurate. Wayne's median length for non-enhanced sentences was longer than Oakland's because of the criminal history weight of sentences that qualified for enhancement.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Section III. Policy Implications

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER FIFTEEN. COMPARING MICHIGAN TO OTHER JURISDICTIONS

A. Introduction

As noted in Chapter One, sentencing schemes in the United States vary enormously. Reforms over the last 45 years have focused on constraining discretion in two ways.

- Determinate or "flat" sentencing sets a single term of years instead of a minimum and maximum sentence. Determinate sentencing eliminates the role of the parole board in deciding when a person will actually be released from prison.
- Structured sentencing is meant to enhance fairness and consistency in judicial decision-making about what sentence should be imposed. There are two common forms.
    - In presumptive sentencing, the Legislature sets narrow penalty ranges based only on the severity of the offense in statute. Judges must impose a sentence within the applicable range unless they find circumstances warranting a departure.
    - Sentencing guidelines, which may be mandatory or advisory, prescribe penalties based on both offense severity and the defendant's prior record. The guidelines may be calculated based on a grid or on a set of worksheets.

Virtually every possible combination of indeterminate versus determinate sentencing and structured versus non-structured sentencing exists in some jurisdiction. While some states apply different schemes depending on the offense, the National Conference of State Legislatures has summarized the primary systems used by each.[1]

- Twelve states use both determinate sentences and structured sentencing, as do the District of Columbia and the federal courts.[2]
- Five states have determinate sentences but do not structure judicial sentencing decisions.[3]
- Fourteen states retained indeterminate sentences but use some form of structured sentencing.[4]
- Nineteen states use both indeterminate sentences and non-structured sentencing.[5]

In this mix, Michigan is among 14 jurisdictions that use a grid style of sentencing guidelines that places criminal history on one axis, offense severity on the other axis and the recommended sentence or sentencing range in a cell where the two axes intersect.[6] We have selected six of these jurisdictions—Kansas, Minnesota, North Carolina, Oregon, Utah and the federal courts—to provide a basis for comparison. While time and space do not permit an exhaustive review of all the comparable jurisdictions, these are representative and allow us to place Michigan's system in a broader context.

We begin with a brief overview of how each of the other jurisdictions' guidelines is structured. Next we apply each jurisdiction's guidelines to a set of hypothetical cases so that we can compare the sentence recommendations that would result. Finally, we compare how these jurisdictions have addressed issues that are common to every system, such as how to score prior record, how to define offense severity, and the basis on which to allow departures from the recommended sentence.

A few points should be noted.

The purpose of scoring the hypothetical cases is to compare the presumptive sentences or sentencing ranges that would result directly from the application of each jurisdiction's guidelines. We do not attempt to speculate about whether a judge would choose to depart from the guidelines recommendation. We also do not speculate on whether a judge would impose sentences to run consecutively, where that is an option. We do indicate when the offense would be subject to a mandatory minimum sentence that would supersede the minimum sentence recommended by the guidelines. To help the reader put the presumptive guidelines sentence in context, we also provide information about how much of the minimum sentence the defendant would be required to serve.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Thus, the charts should not be read as predicting what the actual sentence in a given jurisdiction would be or how much time the hypothetical client would actually serve in prison. Rather, they should be used to compare the allowable sentence lengths and the breadth of judicial discretion available for the hypothetical offense as committed by the hypothetical defendant across guidelines systems.

Every effort has been made to score each hypothetical case accurately.[7] This required identifying the most comparable offense from each jurisdiction's penal code, understanding the nuances of each scoring system, and applying any required statutory enhancements.[8] Given the complexity of the effort, we cannot guarantee that every judgment call made would be universally agreed upon. We are confident, however, that, together, the overviews and the hypothetical cases provide a broadly accurate and fair basis for placing Michigan's guidelines in a national context.

B. Overview of Six Jurisdictions

1. Kansas[9]

The Kansas Sentencing Guidelines took effect in 1993, with the goal to seek equity among like offenders in typical case scenarios.[10] The 17-member Sentencing Commission, which is an independent agency within the executive branch, has 14 staff members.[11] To avoid overcrowding, whenever the prison population equals or exceeds 90% of total capacity, or will do so within two years, the Commission is required to propose guideline modifications for legislative approval.[12]

Kansas abolished parole for all felony crimes except those designated as off-grid—capital murder, first-degree murder, treason, terrorism, and certain sex offenses involving victims under age 14. Actual time served on guidelines sentenced may be reduced by good behavior credits of either 15 or 20%, depending on the offense and when it was committed. Up to 60 days credit for program participation may also be applicable.

Kansas utilizes two grids: one each for drug and non-drug offenses. The drug offense grid has five severity levels; the non-drug offense grid has 10.[13] The intersection of the criminal history and offense severity levels yields a grid block containing three numbers. These indicate, in months, the prescribed aggravated, standard, and mitigated determinate sentences.[14] Judges are advised to select the midpoint or standard term in the usual case and use the upper or lower term to take into account any aggravating and mitigating factors that do not amount to sufficient justification for a departure.[15]

The Kansas penal code no longer sets a single statutory maximum sentence for the crime. Instead, the highest sentence in the applicable grid block is considered the presumptive statutory maximum sentence.

Kansas criminal statutes designate each crime's severity level and categorize it as a person or a non-person crime. For example, KSA 21-5420 defines the crimes of robbery and aggravated robbery. It then states that the former is a severity level 5 person felony and the latter is a severity level 3 person felony.

The nine levels of criminal history, labeled A–I, reflect both the number and nature of prior convictions. For example, Category A is the most serious, for offenders with three or more prior felony convictions designated as person crimes. Presumptive sentence lengths drop dramatically between Categories B (two prior person felonies) and C (one person and one non-person prior felonies).

There is no decay factor for adult convictions; all are scored regardless of age, including expungements. Some juvenile adjudications are not included if the current crime is committed after age 25. Otherwise, prior juvenile adjudications are treated the same as adult priors for criminal history classification.[16] Being on adult probation or parole at the time of the sentencing offense does not affect criminal history scoring, but requires consecutive sentencing unless that would result in manifest injustice.[17]

Presumptive sentences may be subject to "special sentencing rules." Among the most common of these are consecutive sentences if the crime was committed while the defendant was incarcerated, on probation, or parole or on bond for a felony, and presumptive incarceration if a firearm was used in the

RECEIVED by MSC 4/12/2022 1:07:49 PM

commission of a person felony.[18] Incarceration is also presumed if the crime was committed for the benefit of a street gang, but a non-prison sentence is allowed if findings are made about available treatment or the promotion of offender reformation.[19]

Sentence lengths may be enhanced for certain sex offender designations. "Persistent sex offender" status may double the maximum of the presumptive prison term;[20] "aggravated habitual sex offenders" must be sentenced to imprisonment for life without the possibility of parole.[21] "Predatory sex offender" is an aggravating factor that permits an upward departure. By statute, any prior conviction used to establish predatory sex offender status is also included in determining the criminal history category.[22]

Kansas places two notable constraints on sentence length. Upward departures may not exceed twice the maximum presumptive sentence.[23] Consecutive sentences for multiple current convictions, which generally may be imposed in the court's discretion, may not exceed twice the "base sentence." The "base sentence" is the one actually imposed for the primary crime— not the longest sentence in the grid block. Moreover, the non-base sentence is calculated using only the offense severity level of the non-base crime. The criminal history level is assumed to be the lowest one possible.[24]

Compliance with the Kansas guidelines is mandatory. Judges must sentence within the prescribed guideline range, departing only when substantial and compelling circumstances exist.[25] Any fact that would increase the penalty for a crime beyond the statutory maximum other than a prior conviction must be submitted to a jury and proved beyond a reasonable doubt or admitted by the defendant (e.g., as part of a plea agreement.)[26] By statute and case law, appellate courts lack jurisdiction to consider a challenge to a correctly scored presumptive sentence.[27] Departures are subject to review on appeal by either party.[28]

The guidelines contain non-exclusive lists of aggravating and mitigating factors that may justify a departure. These factors generally involve case-specific facts that make the conviction offense more or less serious than the typical crime of its type, such as degree of harm to a victim, the victim's vulnerability, the defendant's role in the offense, and the defendant's physical or mental impairment.[29] Kansas does not provide mitigation consideration in the guidelines scoring for a guilty plea or acceptance of responsibility, but a court may consider assistance provided in the investigation or prosecution of another as a mitigating circumstance. A fact that is an element of the current crime can only be used to justify a departure if the fact makes that aspect of the crime significantly different from the usual criminal conduct of its type.[30]

In 2019, of the incarceration guideline sentences, 50.5% were within the presumptive incarceration range. Of the durational departures, 64.9% were downward and 35.1% were upward. Of those within the presumptive range, 35.4% were standard sentences, 15.6% were aggravated, and 25.7% were mitigated.[31]

2. Minnesota[32]

Minnesota's guidelines, which became effective in 1980, are among the oldest and most thoroughly researched of the state guidelines systems. Among their goals are proportionality, parsimony (the lowest sentence necessary to achieve the purposes of sentencing), and avoidance of prison overcrowding. The state has consistently had the lowest, or one of the lowest, rates of incarceration in the country. The 11-member Minnesota Sentencing Guidelines Commission is an independent agency with a six-person staff.

Minnesota abolished parole and adopted determinate sentencing for all but the few crimes for which parolable life terms are imposed: first-degree murder, treason, and certain sex offenses. Defendants sentenced to prison are incarcerated for two-thirds of the imposed term and spend the remaining one-third on supervised release, although incarceration can be extended for disciplinary violations or refusal to participate in programming.

The guidelines present judges with a presumptive term and a range. The bottom of the range is 15% lower than the presumptive term, and the top of the range is 25% higher. Sentences within this range are not considered departures.

Minnesota utilizes three grids: Standard, Sex Offender, and Drug Offender. On each grid, the offense severity level is on the vertical axis and the criminal history is on the horizontal axis. In

RECEIVED by MSC 4/12/2022 1:07:49 PM

constructing the guidelines, the Sentencing Commission reduced the emphasis previously given to criminal history and made the conviction offense the primary factor in determining punishment.

There are seven levels of criminal history. The criminal history score is the sum of the points assigned for:

- Prior felonies.
- Custody status at the time of the offense.
- Prior misdemeanors and gross misdemeanors.
- Prior juvenile adjudications.

The weight assigned to prior felonies depends on their severity. No weight is assigned to an offense for which a judgment of guilty has not been entered before the current sentencing. Prior adult offenses cannot be counted if 15 years have elapsed since the date of the initial sentence following the prior conviction or, in the case of a prison sentence, the date the sentence expired.

Offense severity is determined by the ranking assigned to every felony offense by the guidelines commission. The ranking applies to the conviction offense, not the charged offense. For attempts and conspiracies, the presumptive sentence is one-half of that shown in the applicable cell.

The presumptive sentence may be subject to a statutory mandatory minimum sentence. There are mandatory minimums for first-, second-, and third-degree murder, certain sex offenses, and driving while impaired. There are also mandatory minimums for possession or use of a weapon while committing an offense by either the defendant or an accomplice. They are one year and a day for a dangerous weapon other than a firearm; three years for a first offense with a firearm; and five years for a second or subsequent offense with a firearm. Notably, the court can depart from the mandatory minimums for weapons on motion of the prosecutor if it finds substantial and compelling reasons to do so.

Presumptive sentences may also be increased if statutory modifiers apply. For instance:

- If a felony assault was motivated by bias, the presumptive sentence must be increased by 25 percent.[33]
- If the offense was committed for the benefit of a gang and the victim was under the age of 18, the presumptive sentence is increased by 24 months if the offense was completed or 12 months if the offense was an attempt or conspiracy. If the victim was 18 or older, the increases are 12 and six months, respectively.[34]
- If the current conviction is for a severe violent offense, the presumptive sentence is increased for prior convictions of severe violent offenses: 12 months are added for one such prior, 18 months for two priors, and 24 months for three or more. A conviction for an offense excluded from criminal history score computation because it was too old under the decay rule does not qualify as a "prior severe violent offense conviction."

Compliance with the Minnesota guidelines is mandatory. Judges may only depart for identifiable substantial and compelling reasons that are subject to review on appeal by either the defense or the prosecution.

The guidelines contain non-exclusive lists of aggravating and mitigating factors that may justify a departure. These factors generally involve case-specific facts that make the conviction offense more or less serious than the typical crime of its type, such as degree of harm to a victim, the victim's vulnerability, the defendant's role in the offense, and the defendant's physical or mental impairment. However, aggravating factors also include aspects of the defendant's prior record (i.e., whether the defendant is being sentenced as a "career offender,"[35] an "engrained offender,"[36] or a "dangerous offender who commits a third violent crime"[37]). The existence of aggravating factors other than prior record must be determined by a jury unless the defendant has waived that right (e.g., as part of a plea bargain). Pursuant to case law, departures cannot be longer than double the presumptive term.

RECEIVED by MSC 4/12/2022 1:07:49 PM

The Guidelines also contain a list of prohibited departure factors: race, sex, employment factors, social factors (education, living arrangements, marital status), and "the defendant's exercise of constitutional rights during the adjudication process." Case law prohibits basing a departure on a factor already considered in determining the recommended sentence. In 2019, the departure rate for prison sentences was 23.2%; of these, 20.3% were mitigated dispositions (i.e., downward departures) and 2.9% were aggravated dispositions. The rate of departures varied substantially depending on the judicial district and offense type.[38]

3. North Carolina[39]

North Carolina adopted its Structured Sentencing System in 1994 and abolished parole for crimes committed from that date on.[40] Parole is retained only for crimes committed prior to that date.[41] The system is intended to produce sentences that are truthful, consistent, rational, and prioritize correctional resources.[42] The 28-member sentencing commission is an independent advisory body with a 10-person staff that operates within the state's judicial branch.[43]

In North Carolina, the judge selects the minimum sentence. Although each offense has a statutory maximum, the maximum to be served by the individual defendant is automatically set at 120% of the minimum. For certain sex offenses, the maximum is then increased by 60 months. The defendant must serve the entire minimum but may accumulate credits for work and program participation, which are deducted from the maximum. That is, while no discretionary decisions are made by a parole board, prisoners essentially work their way down from their maximum release date to release on their minimum.[44] Defendants must also complete a period of either nine or 12 months of post-release supervision, depending on the severity level of the offense.[45]

The North Carolina system uses just two grids—one for felonies and one for misdemeanors.[46] The seriousness of the crime is determined by statute, with every crime being assigned to one of 10 offense classes. These 10 classes are listed on the vertical axis of the grid, although Class A offenses carry death or mandatory life without parole and are not scored under the guidelines. To determine offense severity, the court simply identifies in which of the other nine offense classes the crime of conviction falls.

To determine prior record level, the following factors are considered: the number and severity of prior felonies, certain misdemeanors, similarity of the current offense to any prior offense, and custody status. The number of points defining each level range from 0–1 for Level 1 to 18+ for Level VI.

- Prior record points are assigned to each prior felony conviction based on its offense class. The number of points ranges from 10 for a prior Class A felony down to 1 for a Class A1 or Class 1 misdemeanor.
- One additional point is added if all the elements of the present offense are included in any prior offense, whether or not the prior offense was used in determining the prior record level.
- One additional point is added if the offender was on supervised or unsupervised probation, parole, or post-release supervision, serving an active sentence in jail or prison, or was on escape from a correctional institution while serving a sentence of imprisonment at the time the present offense was committed. This factor must either be found by the jury beyond a reasonable doubt or admitted by the defendant.[47]

All prior adult felony convictions are counted. There are no lapse or decay rules, and no statute of limitations on the use of prior convictions to determine prior record level.[48] Juvenile commitments are not counted.[49] Also not included are certain misdemeanors and traffic offenses (except for death by vehicle).[50]

North Carolina employs four statutory enhancements based on the defendant's criminal history that can increase the punishment beyond what the guidelines recommend. For any of them to apply, the prosecution must have opted to charge the specific enhancement and convicted the defendant of it.

RECEIVED by MSC 4/12/2022 1:07:49 PM

- Habitual felon—three prior felony offenses of any kind increases the offense class by four levels but no higher than Class C.
- Violent habitual felon—two violent prior felonies increases the sentence to life without parole.
- Habitual breaking and entering—one or more prior breaking and entering offenses makes next breaking and entering conviction a Class E felony.
- Armed habitual felon—one or more prior firearm-related felonies requires sentencing as at least Class C with a mandatory minimum of 120 months.

Every cell on the grid contains three ranges: presumptive, aggravated, and mitigated. The court must select the minimum from the presumptive range unless the jury finds or the defendant admits aggravating facts beyond a reasonable doubt,[51] or the defendant proves mitigating factors to the court by a preponderance of the evidence. The court has the discretion to use the aggravated or mitigated range if the facts support the choice. Judges must place the reasons for departing from the presumptive range on the record and either party can appeal the adequacy of the fact-finding.

Factors that can aggravate or mitigate the presumptive sentence are listed in statute, although each list ends with a catch-all category for other factors "reasonably related to the purposes of sentencing.[52] While there is no separate list of prohibited departure factors, the presence or absence of prior convictions is not aggravating or mitigating, as prior record is always considered when determining the prior record level.[53]

The more than 20 permitted aggravating factors include the defendant's role in the offense, victim vulnerability, possession or use of a deadly weapon, serious permanent injury to the victim, exceptional cruelty, interference with arrest, and intent to benefit a gang. Notably, they also include juvenile adjudications for what would be serious felonies and the fact that the defendant does not support his or her family.

The 20 mitigating factors include acceptance of responsibility, assistance in apprehending or prosecuting another felon, minor role in the offense, voluntary participation by a victim older than 16 and characteristics that reduce culpability like age, immaturity, and limited mental capacity.[54] Notably, they also include the defendant's military service, positive employment history, reputation and support in the community, post-arrest entry to drug or alcohol treatment, and the fact that the defendant supports his or her family.

Compliance with the North Carolina guidelines is mandatory. Since aggravation and mitigation are built into the scheme there is no provision for departing above the high end of the aggravated range or below the low end of the mitigated range. In 2018, of the prison sentences imposed for person offenses, 64% were in the presumptive range, 29% were in the mitigated range, and 7% were in the aggravated range.[55]

4. Oregon[56]

The Oregon Criminal Justice Commission, which has a nine-person staff, is responsible for Oregon's sentencing guidelines. The guidelines, which became effective in 1989, apply to all felonies except aggravated murder.[57] They were designed to make sentencing more predictable, proportional and fair, and to take prison capacity into account.

Oregon abolished parole and adopted determinate sentencing for all but the few crimes for which parolable life terms are imposed: first-degree murder, treason, and certain sex offenses. Most defendants serve a specified prison duration that can be shortened by the award of good conduct and program participation credits up to 20%.[58] Time served for certain violent crimes cannot be reduced, including any that have a mandatory minimum sentence.[59]

RECEIVED by MSC 4/12/2022 1:07:49 PM

Oregon has a single sentencing grid with 11 categories of crime seriousness (11 down to 1) and nine levels of criminal history (A–I). The court selects the presumptive determinate term from the range of months in the grid block where the axes intersect. Oregon's ranges are extremely narrow.

Crime seriousness is determined by reference to the Crime Seriousness Scale in the Oregon Administrative Rules.[60] For example, first-degree robbery is classified at crime category 9. The Rules also contain a list of subclassified crimes that apply to offenses that cover a broad range of conduct.[61] For example, burglary has three possible crime categories depending on whether there was weapon use or victim injury, and whether the location was an occupied dwelling. Attempts and solicitations are ranked two categories below the completed offense.

The maximum possible sentence is identified in the statute defining the crime. Penal statutes place each crime in one of four felony classifications: Class A, maximum 20 years; Class B, maximum 10 years; Class C, maximum five years; and Unclassified felonies that include penalties described in the section defining the crime.[62] For example, first-degree robbery is a Class A felony.[63]

The applicable criminal history level depends on a combination of the nature and number of prior adult convictions and juvenile adjudications in the defendant's history at the time of the current sentencing. Adult convictions and juvenile adjudications are treated the same for person felonies but differently for non-person felonies. For instance:

- Category A includes three or more adult convictions or juvenile adjudications for person felonies.
- Category E includes four or more adult convictions for non-person felonies.
- Category G includes four or more adult convictions for Class A misdemeanors, one adult conviction for a non-person felony, or three or more juvenile adjudications for non-person felonies.
- An adult with no person felony prior conviction cannot score above Category E. A juvenile with no adjudication for a person felony cannot score above Category G.

Expunged prior adult convictions and juvenile adjudications are not included, nor are prior findings of "guilty except for insanity."[64] Only Class A misdemeanors are considered. There are no rules on decay or other limitation on the counting of older convictions.

More than two dozen felonies require a mandatory minimum sentence.[65] Defendants serving mandatory minimums are not eligible for release on post-prison supervision or any form of temporary leave from custody, or any reduction based on the minimum sentence for any reason.[66] If the mandatory minimum exceeds the presumptive sentence from the guidelines grid, the mandatory minimum controls.[67]

Compliance with the Oregon guidelines is mandatory. The court must pronounce a sentence within the recommended range unless it states on the record substantial and compelling reasons to impose a departure.[68] Upward departures may not exceed twice the high end of the applicable guidelines range. Departures may be appealed by either the prosecution or the defense. However, the Oregon rules expressly permit plea agreements to stipulated grid blocks and to stipulated sentences, whether or not they fall in the presumptive guidelines range.[69]

The guidelines contain lists of non-exclusive mitigating and aggravating factors that may be considered grounds to depart.[70] These include such matters as the defendant's role in the offense, the victim's status as an aggressor or participant in the offense, the degree of harm to the victim, use of a weapon, and the crime being part of an organized criminal operation. There is no mitigation for a guilty plea, but consideration may be given if the defendant cooperated with the state on the current crime or any other criminal conduct.[71] Several potential grounds for departure are expressly prohibited:

- Refusal to cooperate cannot be deemed an aggravating factor.
- A fact used to trigger a mandatory sentence cannot be used as an aggravating factor.

RECEIVED by MSC 4/12/2022 1:07:49 PM

- Crime facts that are elements of the conviction offense or were used to determine its offense severity level can only be used as an aggravating or mitigating factor if the facts make the case significantly different from typical cases of that type.[72]

Sentences imposed on people already serving a sentence for a previous offense must be served consecutively. When the defendant is being sentenced for two or more unrelated crimes, the court has full discretion to make the sentences concurrent or consecutive. When a defendant is being sentenced for multiple crimes committed in a continuous and uninterrupted course of conduct, consecutive sentencing is permitted, but the court must find that one of several enumerated circumstances justifies that decision.[73] Without a departure, the combined length of consecutive prison terms may not exceed twice the length of the maximum term permitted for the most serious of those crimes.[74]

5. Utah[75]

Utah has a 27-member sentencing commission that is part of the Utah Commission on Criminal and Juvenile Justice (CCJJ). The CCJJ includes other agencies that address juvenile justice, substance abuse and domestic and sexual violence. Included in the Sentencing Commission's mandate is increasing equity in sentencing, modifying the guidelines to reduce recidivism and relating sentencing practices to correctional resources. The Commission has two staff members and receives administrative and financial support from CCJJ. The Commission published its first guidelines in 1998. These replaced earlier guidelines drafted by court and corrections officials.

Utah has retained indeterminate sentencing and release through the Board of Pardons and Parole. Minimum and maximum terms of punishment are established by statute. Felonies are classified into four categories:

- Capital (aggravated murder with predicate crimes).
- First-degree, not less than five years up to life.
- Second-degree, not less than one year nor more than 15 years.
- Third-degree, not to exceed five years.

When the defendant is being sentenced to prison, the court imposes a sentence equal to the range defined by statute. The number of months identified on the guidelines grid is not the sentence. Rather, it is the typical time served for the offense by people with the defendant's criminal history. It serves as a recommendation to the parole board, although the board makes the final decision as to when the defendant will be released. The guidelines are purely advisory and have no binding effect on the courts.

While both the court and the parole board are generally prohibited from deviating from the statutory minimum sentence, there are some exceptions. For instance, downward departures from the statutory minimum for "grievous" sexual offenses are allowed upon a finding of mitigating circumstances. Conversely, there are statutory enhancements that add years to the guidelines term for offenses committed in concert with two or more persons, in relation to a criminal street gang, or for repeat and habitual sex offenders. Prisoners may reduce their incarceration time by completing specified programs. At a minimum, there is a four-month reduction in the total period of incarceration for completion of high-priority programs.[76]

The guidelines apply to all felonies other than capital murder. There are also guidelines for misdemeanors and impaired driving. The felony guidelines have three matrices: sex and kidnapping, criminal homicide, and a general matrix for all other offense types.

The sex and kidnapping matrix has 10 levels of crime category; the other two matrices have seven levels. The levels are determined by the nature and degree of the crime.

RECEIVED by MSC 4/12/2022 1:07:49 PM

The sex and kidnapping matrix has three levels of criminal history; the other two matrices have five levels. Criminal history is scored on the basis of six factors:

- Prior felony convictions.
- Prior Class A misdemeanor convictions.
- Supervision history.
- Prior person or firearm convictions.
- Prior juvenile adjudication within past 10 years.
- Most recent postconviction crime-free gap.

Prior felony convictions include only those for which sentencing occurred before sentencing on the current offense. Current offenses are expressly excluded. If multiple convictions arose from a single prior criminal episode, only one conviction is counted. There is no decay factor for adult convictions. However, most juvenile adjudications of delinquency must have occurred within 10 years of the current conviction date. In addition, defendants have points subtracted from their criminal history for crime-free periods preceding the current offense. The number of negative points increases as the time period gets longer.

The guidelines establish a procedure for documenting departures based upon aggravating or mitigating factors. Because the guidelines are only advisory, there is no standard for the substantive review of departures on appeal. Nonetheless, the Commission has identified factors that are prohibited or should be used with caution. Specifically, the Utah Guidelines state:

> Do not include an aggravating factor if: (1) it is already included as an element of the offense (do not double count) or (2) it is an element of the offense but has not been plead to or otherwise proven beyond a reasonable doubt as required by statute and/or case law (the guidelines are not a means to subvert Constitutional principles).[77]

To avoid double counting, aggravating factors should be used with caution if they have already been considered in a risk assessment. Similarly, factors that reflect socioeconomic status more than risk, such as access to treatment or financial stability, should be considered with caution as should factors specific to sex offenses that may not have a statistical correlation with increased risk to reoffend.[78]

Judges have discretion to decide whether a sentence should be concurrent or consecutive. Sentences for defendants who commit their crime while in prison or parole are presumptively consecutive. Otherwise the factors to be considered are the gravity and circumstances of the offenses, number of victims, and the history, character, and rehabilitative needs of the defendant.[79]

Utah has a unique method of sentencing for multiple offenses. When multiple convictions are to be served concurrently, 10% of the prison stay recommended for each shorter sentence is added to the recommended term for the longest sentence. When multiple convictions are to be served consecutively, the guidelines advise adding 40% of the length of the shortest sentence to the full length of the longest sentence. If there are more than two offenses, the guidelines advise adding 40% of each of the remaining offenses to the full length of the longest sentence.[80]

6. United States[81]

Historical Note

Before the federal sentencing guidelines went into effect in 1987, federal judges imposed indeterminate sentences within broad statutory ranges of punishment, after which the United States Parole Commission would decide when offenders were actually released from prison on parole.[82] The Supreme Court recognized that the broad discretion of sentencing courts and parole officials led to significant sentencing disparities.[83] When Congress enacted the Sentencing Reform Act (SRA) it found: "[e]very day federal

RECEIVED by MSC 4/12/2022 1:07:49 PM

judges mete out an unjustifiably wide range of sentences to offenders with similar histories, convicted of similar crimes, committed under similar circumstances."[84]

To decrease disparities and promote transparency and proportionality in sentencing, Congress created the United States Sentencing Commission, a bipartisan expert agency located in the judicial branch.[85] The Commission is composed of up to seven voting members, including a chair, who are nominated by the President and must be confirmed by the Senate. No more than four Commissioners can be from the same political party, and at least three Commissioners must be federal judges.[86] The Commission maintains about 100 employees. [87]

As the nation's prison population continued to swell, so did criticism that the guidelines had produced unduly harsh sentences for non-violent offenses, particularly those tending to involve non-white and poor defendants at disproportionate rates. Challenges were raised to their mandatory nature.

The guidelines were designed to establish a sentencing range based on the defendant's admissions or the jury's verdict. However, the trial court could then exceed that range based on findings of fact it made about the offense based on a preponderance of the evidence. In *United States v. Booker*, 543 US 220 (2005), the trial court found the defendant had actually possessed a lot more drugs than the amount for which he was convicted. That finding mandated a much higher guidelines range, so the court increased the sentence by almost 10 years above the base range for the crime of conviction.

The *Booker* Court held that judicial findings not subject to proof beyond a reasonable doubt that result in a sentence greater than otherwise mandated by the guidelines violate the Sixth Amendment right to trial by jury. The Court's solution was to make the federal guidelines advisory. Notably, the Court had no problem with allowing the new guidelines range to be based on the "real" offense so long as the sentencing judge was not required to apply it.

Federal judges must now consult the guidelines, but are not required to impose the sentences recommended. The Supreme Court has held that the guidelines are the "lodestone"[88] and "should be the starting point and the initial benchmark"[89] of federal sentencing. District courts must still "give 'respectful consideration' to the Guidelines."[90] Therefore, on appeal, the guidelines ranges (and the degree of departure from guidelines) must still be taken into account.[91]

Initially, sentences within the federal guidelines were presumed to be reasonable.[92] Since the guidelines became advisory, appeals courts review the reasonableness of all sentences, including those that fall within the guidelines range.[93] This reasonableness review applies a "deferential abuse of discretion" standard that does not require extraordinary circumstances to justify a sentence outside of the guidelines range[94] and does not require an appellate court to presume that a within-guidelines sentence is reasonable.[95]

The role of the Sentencing Commission post-*Booker* remains important for its ability to base its recommendations on empirical data and national experience, guided by professional staff.[96]

The Current Federal System

Federal sentences are determinate. Parole has been abolished but prison sentences are followed by a term of supervised release of several years, depending on the offense. The time actually served can be reduced through the award of good conduct credits of up to 54 days per year.[97]

The federal sentencing guidelines are extremely complex. The court must follow the three-step process set forth by the Supreme Court in *Gall v. United States*.[98] First, the court must properly determine the guideline range.[99] Second, it must determine whether to apply any of the guidelines' departure policy statements to adjust the guideline range. Third, it must consider all the factors set forth in 18 USC § 3553(a) as a whole, including whether a variance is warranted.[100]

There is only one sentencing grid, called the Sentencing Table, for all covered offenses. The offense severity levels are on the vertical axis and the criminal history category on the horizontal axis. The presumptive sentence is within the range that falls in the cell where they intersect. Pursuant to statute,

RECEIVED by MSC 4/12/2022 1:07:49 PM

"the maximum of the range shall not exceed the minimum by more than the greater of 25% or six months."[101]

Each federal crime is assigned a base offense level that will increase or decrease as specific offense characteristics and other special rules are applied. To determine a prospective sentence range, the first step is to determine the crime's base offense level number, or severity level, as assigned in the Sentencing Guidelines Manual. There are 43 base offense levels, with the lower numbers for low-level offenses and higher numbers for the more serious offenses.

Once a crime's base offense level is determined, "special offense characteristics" for each crime may apply, increasing the base offense level. For example, the base offense level for second-degree murder is 38.[102] If the offense is determined to be a hate crime (motivated by actual or perceived race, color, religion, national origin, ethnicity, gender, gender identity, disability, or sexual orientation), the level is increased by up to three levels;[103] if a victim was physically restrained in the course of the offense, there will be an increase of two levels.

"Relevant conduct" is "the range of conduct that is relevant to determining the applicable offense level" under the Guidelines Manual.[104] It is a principle that impacts nearly every aspect of guidelines application. It reflects the sentencing guidelines' consideration of characteristics of the offense and the defendant's conduct beyond the count(s) of conviction alone, while also placing limits, specific to the type of offense, on the range of conduct that is appropriately considered. Relevant conduct affects a defendant's offense level, as determined by Chapter Two as well as the role and multiple count adjustments in Chapter Three, the criminal history calculations in Chapter Four, and the adjustments for undischarged terms of imprisonment in Chapter Five.[105]

If a defendant is convicted on multiple counts, the Federal Sentencing Guidelines instruct the judge to determine an appropriate "combined offense level" that takes into account the rules for incremental punishment for multiple offenses. The most serious offense of which the defendant is convicted is used as the starting point for establishing the base offense level, and then the less serious counts are used to adjust the defendant's total sentence upward.[106]

In addition to multiple count adjustments, some cases require consideration of other adjustments that can increase or decrease the defendant's overall sentence; examples include:

- Acceptance of Responsibility—The judge may decrease the offense level from one to three levels if, in the judge's opinion, the offender accepted responsibility for his or her offense.
- Minimal Participant—If the offender was a minimal participant in the offense, the offense level is decreased by four levels.
- Knowledge of Vulnerability—If the offender knew that the victim was unusually vulnerable due to age or physical or mental condition, the offense level is increased by two levels.
- Obstruction of Justice—If the offender obstructed justice, the offense level is increased by two levels.

After the base offense level is adjusted for these considerations, the defendant's criminal history category must be determined. Points are assigned for prior sentences: felony convictions, some misdemeanors, some juvenile adjudications, and the offender's custody status as the time of the offense. The total points determine the applicable criminal history category (I–VI).[107] The court must then determine whether grounds exist to depart from the criminal history category upward or downward if the category substantially under- or over-represents the seriousness of the defendant's criminal history, and whether the offender qualifies as a career offender or armed career criminal.[108] These grounds may include prior sentences not used in scoring the criminal history and prior adult criminal conduct that did not result in conviction.

RECEIVED by MSC 4/12/2022 1:07:49 PM

In every case, relevant conduct includes actions the defendant performed in preparation for the offense, during the offense, and following the offense to avoid detection. Relevant conduct also always includes the defendant's directions to others. In a case of "jointly undertaken criminal activity," the defendant is liable for all acts and omissions of others that were (1) within the scope of the jointly undertaken criminal activity, (2) in furtherance of that criminal activity, and (3) reasonably foreseeable in connection with that criminal activity.[109]

The scoring of prior convictions is subject to decay rules that depend on the length of the prior sentence. Prior adult convictions with sentences longer than 13 months are not counted if more than 15 years have elapsed since sentencing or final release from custody. If the prior sentence was 13 months or less, the lapse time is10 years. If the prior sentence was 13 months or less and the defendant was younger than 18, the lapse time is five years.[110]

There are several enhancements: doubling the mandatory minimum for certain drugs crimes; increasing the criminal history category for career offenders; increasing the criminal history category for armed career criminals and setting a 15-year mandatory minimum; and a three strikes provision for two prior convictions of a serious violent felony resulting in life imprisonment.[111]

Once the base offense level, criminal history category, specific offense characteristics, and relevant adjustments have all been determined, then the sentencing range can be identified according to the Sentencing Table. At that point, the judge must determine whether applying the guidelines is appropriate, or whether it is more appropriate to "depart" from the guidelines and impose an alternate sentence.

A court may depart upward to reflect the actual seriousness of the offense based on conduct underlying a charge dismissed as part of a plea agreement in the case or underlying a potential charge not pursued in the case as part of a plea agreement or for any other reason that did not enter into the determination of the applicable guideline range.[112] A defendant's refusal to assist authorities in the investigation of other persons may not be considered an aggravating sentencing factor; however, a defendant's refusal to assist authorities may be considered in sentencing within the guideline range.[113]

In the federal system, the terms "departure" and "variance" have different meanings and are applied at different stages of the sentencing process. While departures were always contemplated by the guidelines and criteria for applying them were provided, since *Booker* judges may "vary" from the recommended guidelines sentence—based on statutory sentencing goals and principles—even where the rules would not permit a departure. A "variance" outside the guidelines range should occur after consideration of all relevant departure provisions.

Courts have held that variances are not subject to the guidelines analysis for departures. In some situations, a prohibited ground for departure may be a valid basis for a variance. Variances are not subject to notice requirements applicable to departures. A court may grant a departure and a variance in the same sentence (e.g., a departure for substantial assistance and a variance for the defendant's history and characteristics). In sum, the ability to vary preserves district courts' ultimate ability to impose, regardless of the applicable guidelines range, a sentence that it views to be "sufficient, but not greater than necessary" to serve the goals of sentencing.[114]

There is no general prohibition against double counting in the guidelines: using the same factual basis twice, once to form an element of the crime and then to increase a defendant's guidelines sentencing range. The same conduct may determine the base offense level and trigger the cumulative application of enhancements and adjustments unless a specific guideline instructs otherwise.[115]

There are mandatory penalties related to firearms. For example, use of a firearm during a crime of violence carries a mandatory sentence of at least five years,[116] which increases to seven years if the gun was brandished, and 10 if it was discharged. These minimums are consecutive to the substantive crime, such as robbery.

Unless consecutive sentencing is specifically required under applicable statutes (e.g. because the defendant was incarcerated at the time of the current offense), multiple current sentences generally run concurrently. However, a formula is applied that increases the offense severity level of the most serious crime by counting all the current offenses.

RECEIVED by MSC 4/12/2022 1:07:49 PM

In 2020, the rates of compliance, departures and variances for the federal guidelines, nationally and in federal district courts in Michigan, were as follows. [117]

Table 35. 2020 Federal Guidelines Compliance Rates		
	National	Michigan
Within Range	50.4	47.9
Upward Departure	0.4	1.2
Downward Departure	23.0	14.5
Upward Variance	1.8	2.1
Downward Variance	24.5	34.3

C. Hypothetical Case Results

Charts with descriptions of eight hypothetical cases and details about the presumptive minimum sentences that would apply in each of the comparison jurisdictions appear in Appendix D. The following tables summarize key information from each chart.

The tables show the low and high ends of the applicable presumptive range—the "low min" and "high min." In Michigan, Minnesota, Oregon, and the federal system, the grids present a presumptive range within which the court must choose a term (i.e., a minimum term in Michigan, or a maximum term in the other three jurisdictions. Sentences above and below the ends of the range are departures that must be justified by the facts. Kansas has a single presumptive sentence and single mitigated and aggravated sentences. The court is free to select any of these without it being considered a departure. Therefore the mitigated sentence is treated as the "low min," and the aggravated sentence is treated as the "high min." North Carolina has presumptive, mitigated, and aggravated ranges. The mitigated range is for downward departures, and the aggravated range is for upward departures. Therefore the boundaries of the presumptive range are treated as the low and high mins. The width of each presumptive range is shown. Utah has only a single recommended term that is treated as both the low and high mins.

Any routine sentence reductions are applied to produce the "actual low min" and "actual high min," which permits a more realistic comparison of what the recommended minimum sentence actually means. Kansas and the federal guidelines reflect a 15% good time reduction. In Minnesota, the deduction is not traditional good time. Minnesota defendants serve two-thirds of their prison sentences and are then placed on supervised release.

Finally, to provide further context, the tables also show any mandatory minimum sentence that supersedes the guidelines recommendation and any mandatory sentence enhancement that is statutorily required for possession or use of a firearm.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 36. Summary of Chart 1. Second-Degree Murder, Example 1 *[killed man who injured mother during shooting into home, no priors]*											
Location	Range Width		Low Min.	Time Reduc	Actual Low Min.		High Min.	Time Reduc	Actual High Min.	Gun Enhance.	Mand. Min.
Michigan	120		180		180		300/ Life		300/ Life	24	
Kansas	20		166	-25	141		186	-28	158		
Minnesota	106		261	-86	175		367	-121	246	36	
North Carolina	48		192		192		240		240		
Oregon	n/a		Life		Life		Life		Life		300
Utah	n/a		240		240		240		240		180
United States	58		235	-35	200		293	-44	249	60	

Chart 1. This murder case reveals substantial differences among the jurisdictions in both the low and high ends of the ranges. The difference between the actual low in Kansas and the mandatory minimum in Oregon is over 13 years. The actual high minimums differ by as much as 12 years.

Among the actual low-end sentences, Michigan's are the third shortest. However, among the actual high-end recommended sentences, Michigan's and Oregon's are the longest. For Michigan, this is the result of a range that is 10 years wide combined with a lack of any sentence reductions. Thus, in Michigan, the defendant could be required to serve a minimum of 25 years plus two years for possessing the firearm during the felony without any departure from the guidelines range. Because of their longer required firearm enhancements, the results in Minnesota (23.5 years) and the federal courts (25.75 years) are not far off.

Table 37. Summary of Chart 2. Second-Degree Murder, Example 2 *[home invasion, stole jewelry, struck women who found him, she died next day, priors]*											
Location	Range Width		Low Min.	Time Reduc	Actual Low Min.		High Min.	Time Reduc	Actual High Min.	Gun Enhance.	Mand. Min.
Michigan	210		315		315		525/ Life		525/ Life		
Kansas	25		221	-33	188		246	-37	209		
Minnesota	127		312	-103	209		439	-145	294		
North Carolina	63		254		254		317		317		
Oregon	n/a		Life		Life		Life		Life		300
Utah	n/a		252		252		252		252		180
United States	52		210	-32	178		262	-39	223		

Chart 2. This murder case also shows substantial differences among the jurisdictions. The actual low for the U.S. is 11.5 years below the actual low for Michigan. Except for Oregon's life sentence, Michigan has the longest low and actual low minimums. The actual high minimums differ by as much as 26 years. At nearly 45 years or life, Michigan has an actual high min. that is more than twice as long as those of Kansas, Utah, and the United States—66% longer than North Carolina's and 79% longer than Minnesota's. That is, the Michigan defendant can get a minimum sentence that is anywhere from 17 to 26 years longer than the minimums of other guidelines states without departing from the recommended range.

The cause of Michigan's much higher recommended sentences is the 17.5 year width of the range combined with the lack of any sentence reductions. What pushes the hypothetical defendant into such a high range is Michigan's offense variable structure. Under OV 3, he receives 25 points for injury to the

RECEIVED by MSC 4/12/2022 1:07:49 PM

victim preceding her death. Under OV 6, he receives 50 points because the death occurred during a home invasion, even though the defendant pleaded down from felony-murder.

Table 38. Summary of Chart 3. Assault with Intent to Murder, Example 1
[choked wife, struck son, combative with police, priors]

Location	Range Width		Low Min.	Time Reduc	Actual Low Min.		High Min.	Time Reduc	Actual High Min.		Gun Enhance.	Mand. Min.
Michigan	114		171		171		285		285			
Kansas	26		221	-33	188		247	-37	210			
Minnesota	54		135	-45	90		189	-62	127			
North Carolina	63		251		251		314		314			
Oregon	9		121		121		130		130			90
Utah	n/a		108		108		108		108			60
United States	52		210	-32	178		262	-39	223			

Chart 3. The differences among jurisdictions are equally apparent in this assault case. The difference in the low and actual low minimums ranges from 108 months (9 years) in Utah to 251 months (20.9 years) in North Carolina. Among both types of low-end sentences, Michigan is in the middle.

However, among the high-end minimums, Michigan's sentences are exceeded only by North Carolina. At 285 and 314 months, the actual high mins of these two states far exceed the 160-month average of the other five jurisdictions. Once again, Michigan's potential for very long minimum sentences without a departure results from the width of the range.

Table 39. Summary of Chart 4. Assault with Intent to Murder, Example 2
[gang beating of rival gang member, victim paralyzed, defendant 17 years old, juvenile priors]

Location	Range Width		Low Min.	Time Reduc	Actual Low Min.		High Min.	Time Reduc	Actual High Min.		Gun Enhance.	Mand. Min.
Michigan	150		225		225		375/ Life		375/ Life			
Kansas	10		82	-12	70		92	-14	78			
Minnesota	38		118	-39	79		156	-52	104			1 yr. + 1 day
North Carolina	32		125		125		157		157			
Oregon	4		111		111		115		115			90
Utah			96		96		96		96			60
United States	73		292	-44	248		365	-55	310			

Chart 4. The differences among jurisdictions are even greater in this assault case. Between the U.S. low min of 292 and the Kansas low min of 82 are 210 months (17.5 years). For actual low mins, the difference between these two jurisdictions is 178 months (14.8 years). Michigan's low and actual low mins are second in length only to those of the federal guidelines.

At the high and actual high ends of the presumptive ranges, Michigan's recommended minimum sentences exceed the federal guidelines. Without even considering sentence reductions, Michigan exceeds the high ends of the other jurisdictions by from 18 to nearly 24 years. Once again, these extreme differences are possible because of the great width of Michigan's applicable range.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table 40. Summary of Chart 5. First-Degree Criminal Sexual Conduct, Example 1 *[preyed on girlfriend's daughter, multiple counts of CSC<13 and 13–16, no priors]*											
Location	Range Width		Low Min.	Time Reduc	Actual Low Min.		High Min.	Time Reduc	Actual High Min.	Gun Enhance.	Mand. Min.
Michigan	84		126		126		210		210		300 (cts 1-4)
Kansas	n/a		Life		Life		Life		Life		300
Minnesota	35		180*	-60	120*		215*	-71	144*		
North Carolina	48		192		192		240		240		
Oregon	2		58		58		60		60		100
Utah	n/a		192		192		192		192		300
United States	52		210	-32	178		262	-39	223		360
*Includes 25% statutory enhancement for predatory sexual conduct											

Chart 5. This case involving multiple sexual assaults on a child also presents wide variations in the minimum sentences recommended by each jurisdiction's guidelines. Notably, Michigan ranks fifth in the length of both its low- and high-end sentences because the defendant has no prior record and Michigan weights prior record and offense severity equally. However, the point is moot because five of the seven jurisdictions have mandatory minimum sentences that supersede the guidelines recommendation: 30 years for the United States, 25 years for Michigan, Kansas, and Utah, and 8.3 years for Oregon.

Table 41. Summary of Chart 6. First-Degree Criminal Sexual Conduct, Example 2 *[home invasion, repeated assaults on woman woken from sleep, knife, priors]*											
Location	Range Width		Low Min.	Time Reduc	Actual Low Min.		High Min.	Time Reduc	Actual High Min.	Gun Enhance.	Mand. Min.
Michigan	630		270		270		900/ Life		900/ Life		300*
Kansas	61		592	-89	503		653	-98	555		
Minnesota	63		180	-60	120		216	-71	145		
North Carolina	97		386		386		483		483		
Oregon	9		121		121		130		130		100
Utah	n/a		252		252		252		252		Life, no parole
United States	65		262	-39	223		327	-49	278		
*Mandatory with 4th habitual offender enhancement for "serious" crimes and "listed" priors											

Chart 6. Several things about this sexual assault case stand out. The differences among jurisdictions are greater than for any of the other hypotheticals, including the two murder cases, because of the weight given in several jurisdictions to the defendant's prior record. In addition, the high mins in each jurisdiction are generally longer than they are for any of the other cases. The mandatory sentence of life without parole in Utah is particularly noteworthy.

The weight given to criminal history is enormously magnified in Michigan, where the defendant's status as a fourth habitual offender permits the high end of the presumptive range to be multiplied by 100%. This scheme has two consequences. The range is broadened to 62 years, nullifying any notion that the guidelines are structuring the judge's discretion. And the defendant could be sentenced to a 75-year minimum sentence without having to justify a departure. Moreover, whatever sentence within the guidelines is chosen, the defendant is subject to the mandatory 25-minimum. Notably, given the prior

RECEIVED by MSC 4/12/2022 1:07:49 PM

record and the egregious offense facts, the presumptive range would have been 270–450 months even without the enhancement. A sentence at the high end, without any possibility of good time, would still require the defendant to serve 37.5 years.

Table 42. Summary of Chart 7. Armed Robbery, Example 1
[gas station hold-up, co-D went back to kill clerk, priors]

Location	Range Width		Low Min.	Time Reduc	Actual Low Min.		High Min.	Time Reduc	Actual High Min.		Gun Enhance.	Mand. Min.
Michigan	150		225		225		375/ Life		375/ Life		60	
Kansas	22		206	-31	175		228	-34	194			
Minnesota	23		58	-19	39		81	-12	69		60	
North Carolina	17		67		67		84		84			
Oregon	6		66		66		72		72			90
Utah	n/a		120		120		120		120			60
United States	23		92	-14	78		115	-17	98		60	

Chart 7. This armed robbery case also shows substantial variations among the jurisdictions at both the low and high ends of the presumptive ranges. The actual low mins are 6.5 years or less in four jurisdictions; the actual high mins in the same jurisdictions are eight years or less. The ranges are less than two years in every jurisdiction but Michigan. At 12.5 years, the Michigan range is more than six times wider than that of the next widest one. Michigan's actual low min is 50 months longer than the next longest one. Michigan's actual high min is 15 years longer than the next longest one—and that does not include the mandatory five years for the defendant's second conviction of possessing a firearm during the commission of a felony.

These discrepancies have two causes. One is Michigan's placement of armed robbery on the "A" grid, along with assault with intent to murder and first-degree criminal sexual conduct, making the presumptive ranges for all three crimes the same. Other jurisdictions tend to treat robbery as a less severe offense.

Second is the scoring of the offense variables. In Michigan, this defendant receives 100 points under OV 3 for the clerk's death, although the murder charge was dismissed in plea negotiations. He also receives 25 points under OV 1 for his co-defendant's discharge of a firearm and 25 points under OV 13 because this robbery and murder and his prior conviction for unarmed robbery constitute three crimes against a person in a five-year period regardless of whether there was a conviction. Without these 150 points the defendant's recommended sentence would be 81–135 months—still high but much more in line with those of other jurisdictions.

Table 43. Summary of Chart 8. Armed Robbery, Example 2
[2nd of 3 robberies in 9 days, toy gun, no prior record, no injuries]

Location	Range Width		Low Min.	Time Reduc	Actual Low Min.		High Min.	Time Reduc	Actual High Min.		Gun Enhance	Mand. Min.
Michigan	34		51		51		85		85			
Kansas	11		89	-13	76		100	-15	85			
Minnesota	19		50	-17	33		69	-23	46			
North Carolina	17		67		67		84		84			
Oregon	1		39		39		40		40			90
Utah	n/a		72		72		72		72			60
United States	11		46	-7	39		57	-9	48			

RECEIVED by MSC 4/12/2022 1:07:49 PM

Chart 8. In this armed robbery scenario, the absence of both injury to a victim and a prior record for the defendant make the recommended sentences both shorter and more similar across jurisdictions. Nonetheless, there is nearly a five-year difference between the actual low end of the presumptive range for Minnesota and the 90-month mandatory minimum imposed by Oregon. Michigan is closer to the low end. Among the actual high mins, the difference between the extremes is less than four years, with Michigan close to the high end. The width of Michigan's range is still higher than those of other jurisdictions, and the case is still scored on the "A" grid, but since this defendant falls at a relatively low point on that grid there is less room for disparity.

Overall Trends. Despite using structurally similar methods for determining minimum sentences, these guidelines jurisdictions often reach very different conclusions about what sentence is proportional to the facts. The low and high ends of the guidelines ranges differ greatly, and those differences continue when the presumptive sentences are adjusted for routinely awarded reductions. The guidelines vary in the weight they put on the defendant's criminal history, the extent to which they are superseded by mandatory minimum sentences, and whether their recommended sentences are extended by firearm enhancements.

A few trends are common. Sex offenses against children are treated with great harshness. Guidelines sentences are often superseded by statutorily mandated minimums, which are most commonly 25 years. Sex offenses by people with prior convictions for sex offenses also bring extremely long sentences, although the mechanism is more likely to be some form of criminal history enhancement. In the comparison cases, the sexual assaults sometimes brought sentences that were equal to or greater than the murders. On the other hand, recommended minimum sentences for armed robbery tend to be relatively lower.

Michigan is notably different from the other guidelines jurisdictions in several important ways. It has the widest ranges by far, meaning that Michigan judges have much more discretion to impose very long sentences without having to justify a departure from the guidelines recommendation. Michigan requires details about the severity of the offense to be scored that may or may not be used as aggravating factors in other jurisdictions, thereby pushing some defendants into higher ranges. Nonetheless, Michigan does not routinely have the longest sentences at the low ends of the applicable ranges; that is, compared to the other jurisdictions, the starting points of the applicable Michigan ranges are frequently higher than some but lower than others. However, at the high end of the presumptive ranges Michigan's recommended sentences are consistently the longest or one of the longest. Michigan's relatively long guidelines recommendations are compounded by the fact that Michigan allows no sentence reductions for good behavior and it regularly employs every strategy available for lengthening selected sentences, including firearm enhancements, habitual offender enhancements, and mandatory minimum sentences that trump the guidelines recommendations.

D. How Does Michigan's Guidelines System Compare?

The interlocking details of each jurisdiction's guidelines system are so unique that it is difficult to compare specifics. One cannot, for instance, compare the weight given to certain prior record variables by looking at the points awarded because all the point schemes are different. However, there are certain issues that are common to all guidelines systems, and it is possible to broadly compare how the example jurisdictions have resolved them. We saw in the last section that Michigan opted to preserve substantial judicial discretion by having far wider guidelines ranges than other jurisdictions. We will now turn briefly to several other fundamental matters.

1. Commission Oversight

Every jurisdiction except Michigan has a commission that has existed since their guidelines were developed. They vary in the size of their membership, from seven (Oregon and United States) to 27 (Utah) and 28 (North Carolina). They vary in the amount of professional staff they have available, from two in Utah to 100 in the federal system. Apart from these extremes, the average is about nine. They also

RECEIVED by MSC 4/12/2022 1:07:49 PM

vary as to where in state government the commission is located and how much authority it has to amend guidelines on its own.

However, every jurisdiction assumes the need to have a commission that not only develops guidelines in the first instance but also:

- Monitors their implementation.
- Collects, analyzes and publishes data regarding their effectiveness.
- Recommends changes necessary to promote the guidelines' goals.
- Responds to changes in the law and public policy.
- Avoids the inconsistencies that tend to arise from ad hoc legislative amendments.

The lack of a commission or any central body to perform these tasks in Michigan is unique.

2. Enforceability

In four of the jurisdictions we examined—Kansas, Minnesota, North Carolina, and Oregon—compliance with the guidelines is mandatory. In all but North Carolina, the sentencing court can depart from the presumptive range for substantial and compelling reasons. North Carolina has ranges within which aggravated and mitigated sentences can be imposed. Judges can choose to sentence within one of those ranges if adequate facts exist but cannot depart from them. In Kansas, Minnesota, and North Carolina, aggravating facts must have either been admitted by the defendant or found by a jury.

The Utah guidelines have been advisory since their inception. The federal guidelines began as mandatory and were made advisory by the U.S. Supreme Court's 2005 decision in *United States v. Booker*.[118] The Michigan guidelines also began as mandatory and were made advisory by the Michigan Supreme Court's 2015 decision in *People v. Lockridge*.[119] When the guidelines are advisory, compliance per se is not an issue. That is, while a party may be able to appeal a sentence that is not within the presumptive range, the appeal must allege that there is something impermissible about the basis or length of the departure, not simply that the sentence deviated from the guidelines recommendation.

The fact that guidelines are mandatory in no way guarantees that sentences will actually be within the presumptive ranges. We saw in Chapter Seven that when Michigan's guidelines were mandatory, the compliance rate for the life-maximum offenses was 55%. The jurisdictional overviews in Section B of the current chapter show that overall compliance rates were 50% in Kansas, 77% in Minnesota, and 74% in North Carolina. In 2020, the advisory federal guidelines had compliance rates of 50%. Downward departures far exceeded upward departures in every instance.

3. Establishing Offense Seriousness: Grids and Offense Severity Levels

The purpose of having separate grids is to have the starting points of the ranges vary not only by the severity of the offense and the defendant's criminal history but also by differences in the offenses themselves. Three jurisdictions have only one grid for felonies—North Carolina, Oregon, the United States. Kansas uses two grids for drug and non-drug felonies. Minnesota has three grids: standard, sex, and drugs. Utah also has three grids: general, homicide, sex and kidnapping.

The grids have one axis that reflects offense severity. Most of the state grids have from seven to 11 offense severity levels. These levels are determined by statute or administrative rule or sentencing commission directive. They are based simply on the statutory definition of the conviction offense. For example, a murder may be level 1, an armed robbery may be level 3, an unarmed robbery may be level 5, and breaking into an unoccupied building may be level 7. The intersection of the offense severity and prior record levels identifies the applicable guidelines range. In some states, required statutory enhancements may be applied—such as for weapons use or victim injury. Beyond that, no attempt to adjust the offense severity based on the facts of the individual offense is made at this point. Rather,

RECEIVED by MSC 4/12/2022 1:07:49 PM

aggravating and mitigating aspects of the offense are used to support departures once the presumptive range has been established.

The federal guidelines are different. They have 43 "base levels" on a single grid, which can be adjusted up or down depending on the existence of various offense characteristics. First, there must be consideration of certain "special offense characteristics," such as use of a weapon, victim injury, and victim asportation. There can then be further adjustments for "relevant conduct" of various kinds. For example, if the offense was a hate crime, it goes up by three levels. Knowledge of victim vulnerability and obstruction of justice each causes two-level increases. Minimal participation can bring a four-level decrease; acceptance of responsibility may result in a decrease of one to three levels. All of these level shifts require judicial fact-finding that occurs as part of the process of determining the presumptive range. Once the base level is determined, the intersection with the defendant's criminal history score identifies the presumptive range. After that, the court can use additional consideration of aggravating and mitigating factors as a basis to depart.

Michigan's guidelines combine the use of multiple grids and the scoring of offense characteristics in order to determine the presumptive range. As noted in Chapter Six, Michigan uses nine grids. Except for the M2 grid that is used exclusively for second-degree murder, Michigan's grids are not based on the nature of the offense but on the statutory maximum penalty, which is ostensibly a proxy for the seriousness of the crime. Thus armed robbery[120] is on the A grid along with first-degree criminal sexual conduct,[121] because they both carry a maximum penalty of life or any term. Providing false information when applying for a state ID, third offense,[122] and manslaughter[123] are both on the C grid because they carry a maximum of 15 years. Larceny from a person[124] and possession of methamphetamine[125] are on the D grid because their maximum penalty is 10 years.

There are then three, four, or six levels of offense severity on each grid that are determined by scoring points on up to 20 offense variables (OVs). These points are assigned based on a judicial determination of the existence of various aggravating characteristics. Judges are required to score the maximum possible points for every applicable OV. No mitigating characteristics are scored. It is the intersection of the offense severity level and the prior record level on the applicable grid that identifies the presumptive minimum sentence range.

Where other states base a presumptive sentence on the general nature of the crime and then make individualized adjustments, Michigan's system is more like the federal guidelines in that it allows an assessment of individual characteristics to determine the presumptive minimum sentence range. But Michigan goes a step further. It actually considers offense severity twice before identifying the presumptive range. First it uses crime seriousness to select the grid, thereby making the recommended sentences more or less severe. This step is comparable to other states' defining offense severity based on the statutory classification of the crime. Then, Michigan moves defendants into increasingly higher ranges on that grid depending on the required scoring of aggravating factors. This consideration of individual factors is what other states do as an exercise of judicial discretion to depart after the presumptive range has been established. Of course Michigan judges can also choose to depart as well.

4. Defining Prior Conviction

One might imagine that defining what counts as a prior conviction would be a straightforward exercise. In fact it depends on the question: prior to what? If the goal is to punish people more severely who continue to commit crimes after having been punished for prior crimes, the logical definition would be a conviction that was entered prior to the commission of the current offense. However, in the jurisdictions examined here, the approach tends to be more literal. A prior conviction is one that preceded the sentencing on the current offense.

- In Kansas the prior conviction had to have occurred prior to the current sentence regardless of when the prior crime was committed. However, the prior cannot be another count in the current case.

RECEIVED by MSC 4/12/2022 1:07:49 PM

- In Minnesota, a prior conviction is weighed if the judgment of guilty was entered before the current sentencing. If multiple prior convictions arose from a single course of conduct, only the most serious offense is counted unless there were multiple victims. When multiple current convictions from a single course of conduct are being sentenced at the same time, one does not increase the criminal history score of another.
- In North Carolina, a person has a prior conviction when, on the date a criminal judgment is entered, the person being sentenced has been previously convicted of a crime.
- In Oregon, the prior sentence had to have occurred before the current sentence.
- In Utah, the prior sentencing had to have occurred before the sentencing on the current offense. However, there is an express prohibition on counting multiple priors from a single criminal episode or counting additional current offenses.
- The federal guidelines award points for priors based on the length of the sentence imposed. Thus, by definition, the prior sentence had to have occurred before the current sentencing. However, the prior sentence cannot be for conduct that was part of the current offense. Multiple sentences are not counted if they arose from a single offense or were imposed on the same day. Federal judges are permitted to depart upward or downward from the criminal history category they believe to be inadequate.

Michigan approaches the issue from two directions. The scoring instructions for PRVs 1-4 state that a prior adult felony or juvenile adjudication is counted if the conviction or juvenile disposition was entered before the sentencing offense was committed. However PRV 7 undercuts this position by scoring as many as 20 points "if the offender was convicted of multiple felony counts or convicted of a felony after the sentencing offense was committed."[126] That is, points are scored if the defendant is currently being sentenced for multiple offenses arising from the same transaction, a practice other jurisdictions expressly prohibit. Michigan then goes a step further and allows up to 20 points for offenses that not only may not have been committed before the current offense but for which the defendant may not yet have been sentenced. It does not appear that any other jurisdictions contemplate this scenario.

5. Aggravation, Enhancement, and Habitualization

Every jurisdiction employs one or more methods of increasing the presumptive sentence based on specific facts about the defendant or the manner in which the crime was committed. The techniques used, the grounds for increase, and the impact on sentence lengths vary so widely that no exhaustive catalog will be attempted here. However a few general observations will provide perspective.

Other states typically provide non-exclusive lists of factors that might be used to aggravate or mitigate the presumptive sentence. The potential aggravating factors often include offense details that are scored in Michigan's offense variables, such as use of a weapon, injury to a victim, number of victims, extreme physical abuse, knowledge of the victim's vulnerability, gang affiliation, exceptional property loss, and the defendant's role. The critical difference is that whether to consider these details, how much weight to give them, and whether they are offset by any mitigating factors is up to the sentencing court and subject to appellate review. In Kansas, Minnesota, and North Carolina, aggravating facts must have been admitted by the defendant or found by a jury.

In the Michigan scheme, aggravating factors built into the offense variables must be scored at the maximum possible number of points to determine the presumptive sentence in the first instance. Their existence is determined by the court based on a preponderance of the evidence. Additional aggravating or mitigating circumstances can then be considered if Michigan judges choose to depart from the presumptive range. The federal guidelines build both aggravating and mitigating characteristics of the offense into the presumptive sentence by adjustments to the base level. They then allow for consideration of additional factors in departures and variances.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Statutory enhancements are terms of years added to the presumptive sentence for specific circumstances, such as use of a firearm, bias crimes, or gang involvement. The primary such enhancement in Michigan is for possessing a firearm during the commission of a felony. Enhancements may have to be sought by the prosecution before they are applicable but the increase in sentence is then required.

Jurisdictions also commonly use enhancements to increase the presumptive sentence based on specific aspects of the defendant's criminal history. Multiple levels of enhancement for prior sex offenses or prior violent offenses are common and may be very severe. Jurisdictions also vary in the extent to which the nature of prior convictions is used to score the defendant's criminal history in the first instance. For example, the number of points awarded may depend not only on the number of prior convictions but also on whether they were for person or non-person crimes. The federal guidelines increase the criminal history category for "career offenders" and for "armed career criminals" and has a three strikes provision that brings life imprisonment for two prior convictions of a serious violent felony.

Michigan is unique in the breadth of its enhancement for habitual offenders. Criminal history is scored based on the number and severity of prior convictions (as measured by the maximum penalty, not by the nature of the offense). But if the prosecutor chooses to enhance, the high end of the guidelines range can be increased based on the sheer number of prior convictions regardless of their nature or seriousness. Thus, these enhancements are not aimed at repeated sex offenses or repeated violent crimes. A fourth offender enhancement in Michigan may be based on nothing but three prior property crimes and may increase the high end of the presumptive range by 100%. It appears that the closest any other jurisdiction comes is North Carolina, which allows three prior felony offenses of any kind to increase the offense class by four levels but no higher than Class C.

6. "Real Offense" Sentencing

"Real offense" or "relevant conduct" sentencing permits the court to determine the severity of the conviction offense based on all the criminal acts the defendant may have committed. This model goes well beyond consideration of aggravating facts about how the conviction offense was committed. "Real offense" sentencing can increase the punishment for the conviction offense based on charges that were reduced or dismissed as part of a negotiated plea and charges that were not even brought.

Under the federal guidelines, the base level of the conviction offense may be adjusted for "relevant conduct," which is defined as: "all acts and omissions committed, aided, abetted, counseled, commanded, induced, procured, or willfully caused by the defendant . . . that occurred during the commission of the offense of conviction, in preparation for that offense, or in the course of attempting to avoid detection or responsibility for that offense."[127] Federal judges are also given express authority to depart upward based on conduct underlying a charge that was dismissed or not pursued but not used to determine the applicable guidelines range.[128]

Michigan's OV 12 requires the scoring of up to 25 points for felonious criminal acts that occurred within 24 hours of the sentencing offense and that have not and will not result in a separate conviction.

Under both these provisions, if a defendant pleads guilty to robbery, for example, in exchange for the dismissal of charges for assault and resisting arrest, the presumptive sentencing range for the current conviction is increased based on the additional criminal conduct.

Real offense sentencing presents multiple problems.

- The defendant is effectively sentenced for crimes that have not been proven beyond a reasonable doubt, that are determined without various procedural safeguards, and that may be based on double or even triple hearsay.
- The prosecutor can induce sentence increases by choosing to dismiss or not bring charges, and then argue that the conduct underlying those charges must be scored in determining the offense severity level.
- Reliance on charges that were dismissed or not brought to enhance the sentence makes plea bargains illusory.

RECEIVED by MSC 4/12/2022 1:07:49 PM

None of the other jurisdictions we examined utilize real offense sentencing. The Model Penal Code rejects the practice in several provisions.[129]

7. Double-Counting

A final significant policy choice is whether sentences can be enhanced by counting the same conduct more than once. Double-counting, whether of the defendant's prior conduct or an aspect of the conviction offense, raises questions of fairness. Once the fact is accounted for and appropriately scored, the only purpose served by considering it again is to make the sentence harsher (that is, to punish the defendant repeatedly for the same behavior). Absent a policy prohibiting double-counting, there are theoretically no limits to how often the defendant's sentence can be lengthened by repeated consideration of the same fact.

The Model Penal Code provides that a departure "may not be based on any factor necessarily comprehended in the elements of the offenses of which the offender has been convicted, and no finding of fact may be used more than once . . . ."[130] Multiple jurisdictions incorporate these principles either explicitly or implicitly through their lists of factors that may be used to depart from a recommended sentence.

- Utah[131] expressly prohibits using an element of the offense as an aggravating factor.
- Kansas[132] and Oregon[133] both prohibit using a fact that is an element of the offense to justify a departure unless the fact makes the case significantly different from typical cases of that type.
- Similar positions appear in the guidelines of jurisdictions not reviewed in this report. Pennsylvania explicitly prohibits departures based on factors already scored in its guidelines, including the offense of conviction.[134] Delaware also prohibits using an element of the offense as an aggravating factor.[135]

The federal guidelines, on the other hand, contain no prohibition on double-counting. An element of the offense can be used to increase the guidelines range. Departures can be grounded in aggravating details that were already used to increase the base level of the offense.

Notably, Michigan law also prohibits basing a departure on an offense or offender characteristic already taken into account in the guidelines scoring, unless the court finds from the facts contained in the court record—including the presentence investigation report—that the characteristic has been given inadequate or disproportionate weight.[136] Ironically, however, the guidelines themselves do a great deal of double-counting. As described in Chapter Six:

- OVs 1 and 2 double-count the possession or use of a firearm in some circumstances.
- OV 6 scores the intent to kill when it is already an element of the offense and has been considered in determining which grid is applicable.
- OV 10 scores facts that constitute elements of first-degree criminal sexual conduct.
- OV 13 counts the sentencing offense as well as crimes scored in PRV 7.
- OV 19 is scored even though the conduct that threatened the security of a penal institution or interfered with the administration of justice is the basis for the sentencing offense.
- Michigan's use of habitual offender enhancements to broaden the presumptive sentence range allows prior convictions already scored as low severity felonies to increase the penalty by 100%.

In sum, Michigan's actual sentences may or may not be longer than those of other guidelines jurisdictions for comparable defendants convicted of comparable crimes. The sentence imposed depends heavily on whether actual judges choose to depart from their jurisdictions' presumptive ranges. However, the construction of Michigan's guidelines was steered by a series of policy choices that consistently favor harsher penalties. While they often follow the federal example, Michigan's guidelines frequently differ

from the more moderate policies of other states. More than two decades after they were made, the effect of these choices on the proportionality and consistency of Michigan's sentences and the size of its prison population has not been determined. Collectively, these choices are worth re-examination by a sentencing commission that is armed with the benefit of experience and large quantities of data.

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER SIXTEEN. SHOULD THE GUIDELINES BE ENFORCED? THE RESULTS OF REASONABLENESS

The issue of enforcing Michigan's sentencing guidelines actually has three parts. First, could Michigan choose to have mandatory guidelines again? Second, does whether guidelines are mandatory or advisory actually matter? Third, would having mandatory guidelines be desirable? The answer to the first two questions is definitely "Yes." The answer to the third is, "It depends."

A. How Guidelines Can Be Made Mandatory

Allen Alleyne was convicted in federal court of robbery and of using or carrying a firearm in relation to a crime of violence. The federal statute required a minimum sentence of five years for the firearms offense. The mandatory minimum increased to seven years if the defendant brandished the weapon and to 10 years if the defendant fired it. The jury indicated "using or carrying" on its verdict form. However, at sentencing, the judge found brandishing so the sentence necessarily increased by two years.

In *Alleyne v. United States*,[1] the United States Supreme Court held that all facts that are used to increase a mandatory minimum sentence must either be admitted by the defendant or found by a jury beyond a reasonable doubt. Imposing a mandatory minimum sentence based on a critical fact that a judge determined by a preponderance of the evidence violates the Sixth Amendment right to a jury trial.[2]

The Michigan Supreme Court followed *Alleyne* in *People v. Lockridge*.[3] It found that the minimum sentence range of the guidelines is a mandatory minimum because it sets the legally prescribed penalty to be imposed. The majority reasoned that since (1) following the sentencing guidelines was mandatory, and (2) the guidelines increase recommended minimum sentences based on judicial fact-finding, Michigan's guidelines scheme violated the defendant's right to have a jury determine all facts not admitted that increase the minimum sentence. That is, the Michigan guidelines were changed from mandatory to advisory because judicial fact-finding is a required part of scoring offense variables that push defendants into increasingly higher guidelines ranges.[4]

The *Lockridge* Court considered several potential remedies. It rejected the idea that only facts admitted by the defendant or found by the jury should be used to score the offense variables. It also rejected the solution proposed by Court of Appeals Judge Douglas Shapiro, which would make the low end of the guidelines range (i.e., the minimum, advisory while the high end or ceiling continued to be mandatory). Instead, the Court chose to follow the federal example and make the Michigan guidelines entirely advisory. While sentencing courts are still required "to consult the applicable guidelines range and take it into account when imposing sentence," they need not follow the guidelines recommendation.

Lockridge had its detractors. In a vigorous dissent, Justices Markman and Zahra not only disagreed with the majority's reasoning, but also expressed deep concern about once again permitting disparities and excesses the guidelines were meant to reduce. The dissenters described multiple "ironic" consequences of returning the guidelines to advisory status:

- The power of the judiciary is expanded in the name of protecting the right to jury trial.
- Predictability in sentencing is lessened.
- A system designed to limit abuses of state power by curtailing judicial discretion is replaced by one that allows nearly unfettered exercises of discretion.
- The expansion of judicial power comes at the expense of the decision by elected legislators to make the guidelines they set binding.
- In the name of fairness in sentencing, the pursuit of sentencing uniformity and consistency that was "the lodestar" of the guidelines system is replaced by "a rule of deference to the widely disparate judgments of 586 judges."[5]

RECEIVED by MSC 4/12/2022 1:07:49 PM

Notably, in her concurring opinion in *People v. Steanhouse*,[6] then-Justice Larson observed:

> The Legislature can tell us its actual will. In *Lockridge*, this Court decided, as was its duty then, on a remedy that it believed best effectuated the Legislature's intent. If it erred, the Legislature is empowered to install any sentencing scheme that it considers best for the Michigan criminal justice system, limited only by the state and federal constitutions. [fn. omitted] It is certainly better equipped than this Court to weigh the policy options. The ball is in the Legislature's court…

So, could the Michigan Legislature still choose to make sentencing guidelines mandatory? Absolutely. Kansas, Minnesota, North Carolina, Oregon, and Washington all have mandatory schemes. Michigan's guidelines were made advisory solely because judicial fact-finding is a required part of scoring the offense variables. Michigan need only adopt the more typical model, which does not make judicial fact-finding part of the initial guidelines calculation. In this model, offense severity is determined solely by the offense of conviction. Every crime has a severity level assigned to it by the Legislature or sentencing commission. The intersection of offense severity and criminal history levels on the grid produce a guidelines range or presumptive sentence. Only after the guidelines recommendation has been determined may judges engage in judicial fact-finding while considering whether and how much to depart.

Restructuring Michigan's guidelines to shift judicial fact-finding from the guidelines calculation stage to the departure stage may seem hyper-technical, if not outright hypocritical. Either way, judges can find facts, and then increase the defendant's sentence based on those facts. However, there are important differences in the process.

1. Scoring aggravating (but not mitigating) factors as part of the guidelines calculation raises the baseline recommended sentence. If the court feels the resulting sentence is longer than warranted, its only option is to depart downward.
2. Under Michigan's current scheme, judges are required to take selected factors into account and to score them as heavily as the facts allow, regardless of whether the judge believes the factors warrant that much weight under the circumstances. Leaving consideration of aggravating and mitigating factors until after a range has been identified on the basis of the defendant's prior record and the crime of conviction allows the court to decide whether the presumptive sentence alone is proportional. If not, the court can decide why not and assign weight to additional relevant factors accordingly.
3. As we are about to see, mandatory and advisory guidelines result in critically different types of appellate review.

B. Why the Distinction Matters

The most consequential difference between mandatory and advisory guidelines is the standard for reviewing the sentence on appeal. When compliance with sentencing guidelines is mandatory, sentences above or below the recommended range are not permitted unless established criteria for allowing departures are met. The governing rationale is that correctly scored guidelines produce presumptively appropriate sentences, so enforcing the guidelines is a means to that end. When a sentence is appealed, the task of the reviewing court is to assess whether the guidelines have been correctly calculated and applied and whether any departure is justified under the applicable standard.

When guidelines are merely advisory, the rationale is that guidelines can help the decision-maker reach an appropriate sentence and so should be consulted; but compliance with the guidelines is not an end in itself. On appeal, the reviewing court's job is not to determine whether the guidelines have been followed but whether the sentence is permissible on its merits.

RECEIVED by MSC 4/12/2022 1:07:49 PM

1. Review Under the Advisory Judicial Guidelines

When the Michigan Supreme Court decided *People v. Milbourn* in 1990, the sentencing guidelines in effect were the advisory judicial guidelines. As explained in Chapter One, *Milbourn* adopted the principle of proportionality: the most severe punishments should be reserved for the most serious crimes and those offenders who have prior records should be treated more harshly than those who do not. Since any given instance of a particular crime may be more or less serious, the gradation of recommended ranges in the sentencing guidelines can help judges determine the penalty that appropriately reflects a combination of sentencing factors. Departures from the guidelines should be based on reasons that were not already covered by guidelines variables or were not scored points sufficient for the circumstances. Even when some departure is warranted, the extent of the departure is also subject to review for proportionality. Where the departure is disproportionate, the sentence is an abuse of discretion.

Proportionality review under *Milbourn* brought intensive scrutiny to trial court sentencing practices. The results of that scrutiny varied widely. While resentencing was ordered in many cases,[7] a number of extreme upward departures were upheld by the Supreme Court.[8] In 1997, just seven years after it was decided, *Milborn* was so often honored in the breach, at least one Court of Appeals panel concluded the proportionality standard had been "substantially altered."[9]

The Supreme Court has not seemed anxious to address that perception. In *People v. Leon Echols*, the Court of Appeals approved a 75–150 year sentence for second-degree murder imposed despite claims that the length was disproportionate and the trial court had failed to justify the 50-year upward departure. Echols was an 18-year-old defendant who shot and killed a man during an altercation over the purchase of a car. He had one prior conviction for cocaine possession. His recommended sentence under the judicial guidelines was in the range of 10–25 years. The trial court's stated reasons for departing were simply "punishment and [to] protect society."

The defendant did not appeal to the Supreme Court when his sentence was upheld in 1992 but sought postconviction relief nearly 20 years later. The majority denied leave to appeal despite Justice Michael Cavanagh's dissent. Relying on *Milbourn,* Justice Cavanagh found that the extent of the departure was not justified by either the defendant's history or the facts of the offense, which he noted were not worse than other second-degree murders.[10]

2. Review Under the Mandatory Legislative Guidelines

When the mandatory legislative guidelines were enacted, the statute required judges to sentence within the recommended ranges unless they placed on the record substantial and compelling reasons for departing from the guidelines. MCL 769.34 (3). In *People v. Babcock*, [11] the Supreme Court held that substantial and compelling reasons had to:

- Be "objective and verifiable."
- "'[K]eenly' or 'irresistibly' grab our attention."
- Be "of 'considerable worth' in deciding the length of a sentence."
- Exist "only in exceptional cases."

In *People v. Smith,*[12] the Court held that such reasons were needed to justify not only the fact of a departure but also the extent of the particular departure imposed. In *Smith,* the Court considered a case where the guidelines recommended 9–15 years for the sexual abuse of a nine year-old girl, but the trial judge imposed 30–50 years. The Supreme Court found that substantial and compelling reasons for an upward departure existed, but the trial court had failed to justify the extent of the specific departure actually made. *Smith* stated the requirement as: "When departing, the trial court must explain why the sentence imposed is more proportionate than a sentence within the guidelines recommendation would have been."

RECEIVED by MSC 4/12/2022 1:07:49 PM

The Court suggested that one method of justification is to place the departure sentence on the appropriate guidelines grid to see into which range that sentence would fall. One could then see how many additional OV and PRV points it would take for the defendant to reach that range and how the additional points would relate to the defendant's facts.

Responding to a dissenting opinion, the *Smith* majority said that not requiring judges to justify both the fact of a departure and its extent would allow similar defendants to "receive widely divergent departure sentences." It continued: "Any arguably reasonable sentence would be upheld, even if it were not proportionate to the offense and the offender. A lack of meaningful review would inevitably encourage idiosyncratic sentencing."

Justice Markman, concurring in *Smith*, emphasized the Legislature's intent that judicial discretion be constrained in order to eliminate wide disparities. He said the question on review "is not whether the sentence is 'reasonable' (cite omitted) but whether it is *lawful*, i.e. in compliance with both the substance and the procedure of the guidelines."

3. Review Under the Advisory Legislative Guidelines

Since the Supreme Court's 2015 decision in *People v. Lockridge*, Michigan's legislative guidelines are now advisory. The "substantial and compelling" standard for departures has been abolished. An appellate court just needs to find that the sentence was "reasonable."[13]

With trial courts no longer bound by the guidelines, the inevitable question is: how does an appellate court determine whether a sentence is "reasonable?" The short answer is: by returning to *Milbourn*. In *People v. Steanhouse*,[14] the Supreme Court said the proper standard upon appellate review is whether a "departure sentence is so unreasonable as to constitute an abuse of the trial court's discretion." The Court observed that when *Milbourn* was decided, the then-existing judicial guidelines were advisory. It described the relationship between the principle of proportionality and those advisory guidelines: "[T]he key test is whether the sentence is proportionate to the seriousness of the matter, not whether it departs from or adheres to the guidelines' recommended range."

While *Steanhouse* repeated its directive from *Lockridge*—that the guidelines "remain a highly relevant consideration in a trial court's exercise of sentencing discretion"—it affirmed the "key test" of *Milbourn.* The Court amplified the point by stressing that no presumption of unreasonableness arises from sentences outside the guidelines range. That is, while sentences within the guidelines are presumed to be proportional, departure sentences are not presumed to be disproportionate.

The Court rejected a rule that would require "extraordinary" circumstances to justify a sentence outside the Guidelines range and further rejected reliance on rigid mathematical formulas to justify the extent of a departure. It also expressly disavowed dicta in *Milbourn* that said departures should "alert the appellate court to the possibility of a misclassification of the seriousness of a given crime by a given offender and a misuse of the legislative sentencing scheme." The Court wanted to avoid suggesting that the guidelines should almost always control and that departures are presumptively unreasonable.

Over time, caselaw has identified various grounds for upholding guidelines departures under *Milbourn*:

1. The guidelines' failure to accurately reflect the seriousness of the crime.
2. Factors not considered by the guidelines.
3. Factors considered by the guidelines but given inadequate weight.
4. Any prior relationship between the victim and the offender.
5. The defendant's misconduct while in custody.
6. The defendant's expressions of remorse.
7. The defendant's potential for rehabilitation.[15]

The subjectivity of most of these grounds is apparent. On what basis does a sentencing judge decide that the guidelines do not reflect the seriousness of a particular offense or that they give inadequate weight to

RECEIVED by MSC 4/12/2022 1:07:49 PM

particular offense variables or that the defendant lacks the potential for rehabilitation? On what basis does a reviewing court decide that a judge who relies on one of these grounds for departure has abused his or her discretion? At what point do judicial decisions about the "adequacy" of the guidelines effectively overrule the legislative determination of what sentence is appropriate?[16]

In *People v. Dixon-Bey*,[17] the majority sought to address these questions by requiring trial judges to explain why a departure sentence was more proportionate than a sentence within the guidelines would have been. The defendant was charged with first-degree murder for stabbing her boyfriend to death but convicted of second-degree murder by a jury. The defendant had no prior record and the guidelines range was 12 to 20 years. The trial court imposed a sentence of 35–70 years because it believed, the jury's verdict notwithstanding, that the defendant had acted with premeditation. She stabbed the victim twice in the chest, had previously stabbed the victim in the hand, had disposed of the knife, and had once told a witness that she could get away with stabbing the victim in the chest by claiming self-defense. However, the court had no evidence of premeditation that had not been available to the jury.

The majority scrupulously measured the facts against how the applicable offense variables were scored or could have been scored. It noted that the trial court had not explained why the guidelines scoring was insufficient. It observed that non-guidelines factors mentioned by the court were either not unique to the defendant or were otherwise relevant to a proportionality determination. Thus, the majority found no basis for concluding "that the trial court's 15-year upward departure sentence was more reasonable and proportionate than a sentence within the recommended guidelines range would have been." *Dixon-Bey* was a 2-1 decision.

Dissenting Judge Boonstra relied on language from *Babcock* that said when there is more than one reasonable and principled outcome, it is not an abuse of discretion for the sentencing court to select either one. He found that the trial judge had heard all the evidence at trial, listened to arguments from both sides at sentencing, and considered the guidelines. Thus, he engaged in a reasoned process that produced a reasoned decision. The abuse of discretion standard gives deference to the trial judge's experience in sentencing and familiarity with the facts. The majority responded to the latter point by noting that giving trial courts that much deference fails to address unjustified disparity and that reference to the guidelines is most important in life-maximum cases where the legislative penalty is so open-ended.

The Michigan Supreme Court considered whether it wanted to grant leave to appeal and review the Court of Appeals decision in *Dixon-Bey*. However, the Court ultimately chose to let the Court of Appeals decision stand.

The substantial and compelling standard assumes that the sentencing guidelines account for the factors commonly associated with the crime of conviction and that the facts have to be unique and exceptional to justify departing from them. The reasonableness standard assumes that a variety of sentences can be justified and that any of them are permissible as long as they are adequately explained. Reasonableness is thus ultimately in the eye of the beholder.

4. Application of the Reasonableness Standard after *Dixon-Bey*

The inconsistencies that result from the reasonableness standard can be seen by comparing the results on appeal in similar cases.[18]

The defendants in *People v. Harper*[19] and *People v. Calloway*[20] both had long prior records and possessed small amounts of drugs.

The 54 year-old defendant in *Harper* was convicted of possessing less than 25 grams of cocaine for having a trace amount of the substance in a crack pipe. He had eight prior low severity felony convictions for such behavior as theft, lying to police indecent exposure and domestic violence and 22 misdemeanors. He also had a history of substance abuse and failure on probation. Because this was his second controlled substance offense, the guidelines range was 4–34 months. Based on this prior record, the trial court departed upward by 14 months and sentenced the defendant to four to eight years in prison.

RECEIVED by MSC 4/12/2022 1:07:49 PM

The Court of Appeals said the departure was "carefully selected to fit this particular offender and this particular offense, being defendant's ninth felony conviction overall." Although the panel recognized that the offense "seems rather trivial standing alone," it could not conclude that the trial court abused its discretion.

The 29-year-old defendant in *Calloway*, was convicted of possessing less than 50 grams of drugs and was sentenced as a third habitual offender. He had 11 prior low severity felony convictions involving drugs, theft, resisting arrest, and fleeing from police. He also had 11 misdemeanors. Because the trial court felt the guidelines did not adequately account for the defendant's prior record, it imposed a minimum sentence of eight years—a departure of more than five years. The Court of Appeals found that, while the inadequacy of guidelines scoring can be a reason to depart, the sentencing judge gave no rationale for the extent of the departure. It ordered a resentencing that resulted in a new minimum of 34 months.

***People v. Nelson*[21] and *People v. Boak*[22] both involved child molestation.**

The defendant in *Nelson* was convicted by a jury of second-degree criminal sexual conduct with a 15-year-old nephew when the defendant was 23 and fourth-degree CSC with the nephew's 13-year-old friend when the defendant was 20. Both incidents involved the defendant having the boy watch pornography and then reaching into the boy's pants and grabbing his penis. The two incidents occurred three years apart. The trial court imposed a minimum sentence of nine years, which was three years above the high end of the guidelines range (as originally calculated before a scoring error was corrected on appeal).

The trial judge based the departure on the defendant's serial preying, his betrayal of trust, and the impact of the offenses on the victims and their families. The Court of Appeals found that two incidents in three years do not constitute a pattern of child victimization, the statutory definition of the crime encompasses the victims' age and family relationship, and the psychological impact on the victims and the defendant's predatory behavior were scored by the guidelines. The trial court had not explained why the guidelines scores were inadequate or why the extent of the departure was warranted.

The defendant in *Boak* pleaded guilty to producing and to possession and distribution of child sexually abusive material. In one incident, she removed the diaper from a sleeping two-year-old and held the child's vagina open while she took a photograph. In a separate incident, she forwarded to co-defendant Thelen a photo of a nine-year-old girl with co-defendant Plowman's erect penis posed near the child. The 54-year-old defendant had no prior convictions. She explained at sentencing that her behavior changed when she started dating co-defendant Thelen.

The trial court imposed a minimum sentence of 12 years, an upward departure of over four years. The trial judge characterized the defendant's conduct as "heinous" and said it "turns my stomach." The court expressed concern about public safety and the need to deter others. It expressed skepticism about the accuracy of a risk assessment instrument that said the defendant was at low risk for re-offending and questioned whether rehabilitation was achievable. It also said the defendant was a willing participant who did not offer an apology until sentencing and then did not address the victims directly. It concluded that the guidelines were inadequate and that a sentence within them would not be reasonable or just. It had sentenced co-defendant Plowman to a minimum of eight years and co-defendant Thelen to 14 months.

The Court of Appeals majority quoted and then summarized the trial court's reasoning but did not compare it to the guidelines scoring. The concurring opinion simply concluded that the record supported both the trial court's decision to depart and the extent of the departure.

The dissent found the trial court's reasons insufficient to establish that the sentence was proportional. It noted that nothing made this crime more heinous than others of its type; in addition the two-year-old was asleep and had no memory of the incident. There was no explanation of why the assessment finding the defendant was at low risk for re-offending should not be trusted or why the defendant's conduct and history made her different from defendants sentenced within the guidelines. The dissent noted that the possession and distribution convictions were scored twice in the guidelines for the

RECEIVED by MSC 4/12/2022 1:07:49 PM

production conviction, resulting in a doubling of the guidelines range for the most serious offense from 24–40 months to 57–95 months. The dissent also quoted the defendant's lengthy and heartfelt statement at sentencing in full.

The defendants in *People v. Mead*[23] and *People v. Tucker*[24] both had long histories of low severity felonies.

Defendant Mead was convicted of Operating While Intoxicated, 3rd offense and was sentenced as a fourth habitual offender. The trial court departed upward by 14 months to impose a sentence of 3–15 years. The reason given for the departure was the defendant's six prior OWIs, his failures in rehabilitation programs, and public safety. In a 2-1 decision, the Court of Appeals majority analyzed the relationship between these reasons and the sentencing guidelines. It found that the prior record had been accounted for both by prior record scoring and the habitual enhancement. It noted that the program failure and recidivism were inter-related. It concluded that the trial court had not explained why the guidelines were inadequate or why the sentence imposed was more proportionate, so it remanded for resentencing.

On the same day that *Mead* was decided, a different Court of Appeals panel issued a unanimous opinion upholding the upward departure in *Tucker*. The 58-year-old defendant was convicted of breaking and entering into a business. Because he was sentenced as a fourth habitual offender, his sentencing guidelines range was 19–76 months. The trial court departed by 32 months and imposed a sentence of nine to15 years. The departure was based on the defendant's extensive criminal history, which included seven felony convictions—primarily for property offenses—and for fleeing from and resisting police, 18 misdemeanors, and multiple probation and parole failures. In upholding the departure, the Court of Appeals found: "Thus, factually, defendant is incorrect when he states that the trial court relied solely on his past criminal history as already contemplated by the guidelines. The trial court clearly also relied on defendant's propensity for recidivism."

The motivation for departures is often judicial dissatisfaction with the actions of other criminal justice actors.

In *People v. Scott,*[25] the defendant broke into the homes of people who were away attending funerals. While he was on bond for a break-in in County A, he committed another one in County B. While on probation in both counties, he committed four more in County C. The County C trial judge found that the prior record and offense variables that scored the other break-ins were inadequate, that the other counties had given the defendant a break, and that he had taken advantage of people in a time of grief. The Court of Appeals found these reasons justified a departure of nearly five years from a range that was already broadened by a habitual third enhancement and upheld the sentence of 19–30 years.

In *People v. Odom*,[26] the defendant was convicted of robbing a payday lender while he was on parole. His prior record made him eligible for enhancement as a fourth offender but the prosecutor had failed to comply with the statutory notice requirements, so the guidelines range could not be broadened. The trial court concluded that the guidelines did not accurately reflect the defendant's recidivism and departed upward by 12.5 years. In finding the 30-year minimum sentence was proportional, the Court of Appeals observed that the defendant's conduct put him in the same position as other repeat offenders notwithstanding the notice error and that had the enhancement been applied the minimum sentence could have been up to 35 years. That is, proportionality was defined by comparison to eligible defendants who received enhancements, as opposed to those who did not. Thus, *Odom* stands for the proposition that no matter why the prosecutor did not seek a habitual offender enhancement, it is reasonable for the sentencing court to achieve the same result through departure.

In *People v. Schurz*,[27] the trial court was apparently dissatisfied with both the prosecution's charging decision and the results of the guidelines scoring. The defendant shared heroin that he had purchased with his best friend from childhood. The friend died of an overdose. The prosecution initially charged the defendant with delivery of a controlled substance causing death, but the defendant pleaded guilty to delivery of heroin under 50 grams. The defendant was scored 100 points under OV 3 for a death occurring during the commission of a crime when the conviction offense was not homicide. He

RECEIVED by MSC 4/12/2022 1:07:49 PM

challenged this scoring because he had injected the same drug without getting sick and the autopsy showed that the victim had seven different drugs in his system, including an extremely large quantity of fentanyl. However, the trial judge found, by a preponderance of the evidence, that the heroin the defendant provided was a cause of the victim's death.

The scoring of OV 3 resulted in a guidelines range of 19–38 months, which was doubled to 38–76 months because this was the defendant's second drug offense. The trial court then departed by 3.7 years, imposing a sentence of 10–40 years. It relied on the defendant's delay in calling 911 and failure to be wholly forthcoming with first responders, the fact that defendant had provided the drugs that caused his best friend's death, and the view that defendant had failed to take responsibility when he said that this was supposed to have been just another instance of casual use as had occurred many times before. The Court of Appeals found these reasons adequately articulated a rationale for the extent of the departure.

Most common is judicial disagreement with a jury's verdict. In *People v. Latimer*,[28] the defendant was acquitted by a jury of assault with intent to murder but convicted of assault with intent to commit great bodily harm and armed robbery. The 21 year-old defendant had a prior record that consisted of 16 misdemeanors. Because they were mostly traffic offenses, only two were scorable in the guidelines. The trial court departed by over 13 years to impose a 32-year minimum sentence for the robbery conviction. Its basis was the brutality of the attack, the defendant's insistence on claiming self-defense, his failure to show remorse, two misconduct citations while in custody, and his prior record. In upholding the sentence, the Court of Appeals majority noted that since the misdemeanors were not scorable they were aspects of his criminal history not counted by the guidelines that justified a departure. It did no analysis of the extent of the departure and concluded: defendant's argument in this case that his "conduct is entirely taken into consideration in the offense variables" is simply not of critical importance to whether or not his sentence is proportionate.[29]

The data analyzed in this report show that proportionality and consistency were elusive goals even when the guidelines were mandatory. These case examples suggest that making the guidelines advisory have undermined those goals further. The data that would systematically test the impact of *Lockridge* on sentence lengths and disparity have not yet been examined. In the short term, it may be that judges who are used to sentencing within the guidelines continue to do so. However, as new judges take the bench, and as judges who chafed under the restrictions of the guidelines feel less constrained, shifts in sentence length may become more pronounced. Anecdotal evidence suggests that lawyers may be having more success in securing downward departures for sympathetic clients, while judges feel freer to impose lengthy upward departures based on their own reactions to the crime or the defendant.

C. Are Mandatory Guidelines Desirable?

The answer depends on the guidelines being enforced. Shifting the timing of judicial fact-finding does not address other problems that have been discussed in this report, such as the breadth of the guidelines ranges, the length of the minimum-minimums, and the nature of the factors used to aggravate the sentence. Mandating guidelines that promote the inconsistent use of habitual offender enhancements, the double-counting of elements of the offense, and departures based on subjective characteristics like attitude and lack of remorse would not be an improvement.

The definition of substantial and compelling reasons employed when the legislative guidelines were mandatory could avert some of these concerns. After all, it required that reasons for departure be objective and verifiable and that they exist only in exceptional cases. But that definition was developed in the context of guidelines that already scored numerous offense details. It would be incumbent on a new sentencing commission to apply that definition to offense severity levels based solely on the offense of conviction and to carefully define what aggravating factors are acceptable. Restructuring the guidelines to allow them to be mandatory is only a first step. Improving them so that compliance is worth mandating comes next.

Mandatory guidelines can work. The experience of other states is instructive. Minnesota's success at combining high judicial compliance rates with relatively low incarceration rates provides an especially

promising example. Achieving similar success in Michigan would require not just making the guidelines mandatory but rethinking when, why, and which departures should be allowed.

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

CHAPTER SEVENTEEN. RECOMMENDATIONS FOR CHANGE

This report has viewed Michigan's sentencing guidelines from multiple perspectives. It has reviewed their history, scrutinized their structure, presented extensive empirical research on their impact, compared them to the guidelines of other jurisdictions, and provided the views of respected commentators. All the analysis leads to a common set of conclusions and supports the following recommendations.

Establish a commission with the mandate and scope of authority of the former Criminal Justice Policy Commission.

Make compliance with the guidelines recommendations mandatory. Allow departures only for substantial and compelling reasons grounded in facts that were found by a jury or admitted by the defendant.

Make fundamental structural changes.

1. Base offense severity scoring that determines presumptive guidelines solely on the crime of conviction. Reserve judicial fact-finding for aggravating and mitigating circumstances that support departures from the guidelines recommendation. Such a system:

    a. Establishes a baseline sentence for the crime of conviction before individualizing punishment based on a wide variety of permutations.
    b. Does not force judges to score the harshest options for each variable regardless of whether they think doing so is appropriate given the facts. Instead, it gives judges the discretion to consider individualized facts in a more nuanced way.
    c. Does not encourage prosecutors to increase the recommended minimum sentence by choosing charges or alleging facts that carry maximum points.
    d. Simplifies guidelines scoring by not requiring complex decision-making about the defendant's guilt of conduct not found by a jury or admitted by the defendant.
    e. Allows the Legislature to return to mandatory guidelines because calculation of presumptive sentences is not based on judicial fact-finding.

2. Weight prior criminal history less heavily than offense severity to avoid imposing punishment that is disproportionate to the harm caused, make better use of scarce prison resources, and reduce racial disparity.

3. Reduce the starting points of cells (the minimum-minimums) on the M2 and A grids to reduce sentence lengths.

4. Narrow the guidelines ranges to reduce disparity and sentence length.

5. Eliminate the use of habitual offender status as a basis for broadening the ranges. This avoids:

    a. Counting prior record twice in selected cases when prior record is already scored in guidelines
    b. Double-counting of specific priors,
    c. Enhancement based on priors that cannot be scored under guidelines,
    d. Undue reliance on priors that may be poor predictors of future risk,
    e. Disparity based on discrepancies among prosecutors' habitual offender charging policies.

RECEIVED by MSC 4/12/2022 1:07:49 PM

6. Redesign grids to:
    a. Reduce their number
    b. Reduce or eliminate redundant cells
    c. Cluster offenses based on their similarity in nature, not just their maximum sentences

7. Change the lowest PRV level to 0–5 points so that two prior misdemeanors or two low severity juvenile adjudications are Level A. Adjust points for other PRV levels accordingly.

8. Limit the allowable length of departures.

9. Limit the aggregate length of consecutive sentences.

Adopt basic principles.

1. Eliminate "real offense" sentencing. In defining offense, consider only facts that are related to the conviction offense.

2. Prohibit using elements of the offense to increase the calculation of offense severity.

3. Prohibit counting a fact more than once unless the fact makes that aspect of the crime significantly different from the usual criminal conduct of its type.

Enact additional recommendations for increasing proportionality and reducing disparity.

1. Move armed robbery to the B grid to better reflect actual sentences and the high rate of downward departures.

2. Move unarmed carjacking to the C grid, where unarmed robbery resides.

3. Do not score multiple prior convictions that arose from the same transaction.

4. Require sentences for similarly culpable co-defendants to be proportional to each other.

5. If a plea agreement is to a sentence "at the bottom of the guidelines range", require that the range be specified during the plea proceeding and that plea withdrawal be allowed if the range changes.

6. Treat felony convictions of people who were 17 at offense as juvenile adjudications.

Recommendations regarding specific prior record and offense variables under the current scheme.

<u>Prior Record Variables</u>

PRV 1 & 2 (High and low severity felonies) – Change breakpoint between high and low severity felonies; define high severity as M2, A and B grid offenses.

PRV 3 & 4 (High and low severity juvenile adjudications) – Change breakpoint between high and low severity juvenile adjudications, as described above. Reduce points for juvenile adjudications. Treat adult convictions by waived juveniles as juvenile adjudications.

PRV 5 (Misdemeanors) – Eliminate scoring of misdemeanors with maximum sentences of 93 days or fewer.

PRV 6 (Relationship to criminal justice system) – Do not score if conduct is basis for sentencing offense and/or will result in mandatory consecutive sentence.

RECEIVED by MSC 4/12/2022 1:07:49 PM

PRV 7 (Subsequent or concurrent felony convictions) – Eliminate. By definition convictions scored are not "priors," and they are often scored in OV 13.

Offense Variables

OV 1 (Aggravated use of a weapon) – Do not score if defendant was also convicted of felony firearm. Only score for unarmed accomplice if they had actual knowledge that principal was armed.

OV 2 (Lethal potential of weapon) – Only score 15 points to avoid double-counting conduct scored in OV 1 and, potentially, felony firearm.

OV 3 (Physical injury to victim) – Do not score 100 points if the accomplice did not have the intent to kill. Do not score if the accomplice was convicted of assault with intent to murder because that is scoring an element of the offense. Do not score if the defendant was charged with but not convicted of homicide because that is penalizing the defendant for the "real" offense.

OV 4 & 5 (Psychological injury to victim or family member) – Require professional assessment of the need for treatment.

OV 6 (Intent to kill or injure) – Eliminate. Scores intent element of offense.

OV 7 (Aggravated physical abuse) – Define "conduct designed to substantially increase the defendant's fear" to be the functional equivalent of sadism/torture/excessive brutality. Alternatively, have a separate category with lower points.

OV 8 (Asportation or captivity) – Create an exception for minor or incidental movement.

OV 9 (Number of victims) – Reduce two victims to five points. Do not include responding public safety personnel in counting the number of victims.

OV 10 (Exploitation of vulnerable victim) – Do not score if the same conduct is the basis for the conviction, e.g. a CSC offense under 750.520 b, c, or d

OV 11 (Criminal sexual penetration) – Eliminate. This is based on "real offense" sentencing.

OV 12 (Contemporaneous felonious criminal acts) – Eliminate. This is based on "real offense" sentencing.

OV 13 (Continuing pattern of criminal behavior) – Eliminate. This double counts convictions already scored *and* is based on "real offense" sentencing.

OV 14 (Offender's role) – Add definition of "leader."

OV 19 (Threat to security of penal institution or court, interference with administration of justice, etc.) – Do not score if conduct was element of offense (e.g., providing contraband to or possession of contraband by a prisoner or jail inmate, witness intimidation or perjury) or for routine attempts to avoid prosecution, such as leaving the scene or not cooperating with police.

RECEIVED by MSC 4/12/2022 1:07:49 PM

CONCLUSION

When the Michigan Felony Sentencing Project issued its report in 1979, it identified three primary concerns about sentencing in Michigan: disparity unrelated to offense severity or to the defendant's criminal history, lack of accountability, and the diffusion of authority over the actual sentence to be served among prosecutors, judges, and parole boards. Unfortunately, despite the development of three sets of sentencing guidelines and the temporary existence of two sentencing commissions, the same concerns remain.

- Sentences continue to vary widely depending on the county of conviction, the identity of the sentencing judge, and whether the defendant was willing to plead guilty. While there is minimal evidence of direct racial bias, non-whites continue to receive disproportionally longer sentences in large part because the guidelines place such heavy weight on prior criminal history.
- Appellate review of sentences for "reasonableness" produces very different results depending on the composition of the appellate panel. Some of the most extreme sentences are reversed, but sentencing judges are largely free to impose whatever departures from the guidelines they choose to rationalize.
- Prosecutors exercise an out-sized influence over sentence length through the ability to add habitual offender enhancements. Enormous differences in the extent to which county prosecutors choose to apply enhancements further add to disparity in sentences for similarly situated defendants.

With the size of the Michigan prison population and the proportion of people serving minimum sentences of more than 10 years having each more than doubled, there is a fourth major concern: the enormous growth in sentence lengths and the pressure that creates on prison resources. The increase in length of time served by Michigan prisoners is compounded by multiple non-guidelines changes, including the elimination of disciplinary credits, the increased use of consecutive sentences and mandatory minimums, and fluctuations in parole practices. However, consistent increases in minimum sentences for life-maximum offenses not only lengthen time served in themselves but also exacerbate the consequences of other "get tough" policies.

The enabling legislation set proportionality, reduction of disparity, and consideration of correctional resources as goals of the sentencing guidelines. The MFSP report warned that broad ranges of sentencing discretion generate disparity and that disparity is inevitably greatest among the life-maximum offenses because that is where the range of discretion available is broadest. Nonetheless, as enacted, the guidelines preserve broad judicial discretion and use multiple mechanisms to encourage extremely punitive sentences for the life-maximum offenses. In many cases, sentencing judges reject this encouragement and impose minimum sentences below the guidelines recommendations. In other cases, judges decide that the recommendation is not harsh enough and choose to depart above the guidelines range. In any event, the high rate of departures suggests judicial dissatisfaction with the guidelines. At a minimum, even when compliance with the guidelines was mandatory, judges appear to have often treated them as simply a starting point for their own decision-making.

Today, Michigan has—in a sense—the worst of both worlds. Unlike other guidelines states, judges must calculate recommended minimum sentences based on guidelines that require them to score every possible aggravating factor, double-count a number of prior record and offense factors, and assign very high weights to selected variables whether or not the judge thinks that weight is warranted by the facts. If followed, the often very punitive results are considered presumptively reasonable. But the guidelines recommendation also creates a baseline from which judges can choose to depart upward for reasons as subjective as their personal feeling that the guidelines scoring did not adequately reflect the

RECEIVED by MSC 4/12/2022 1:07:49 PM

gravity of the crime or the defendant's criminal history. The modest scope of appellate review then permits many of these departures to stand.

Also, unlike other states that use sentencing guidelines, Michigan has no commission to analyze their usage, assess their impact, and recommend changes. Michigan's guidelines have been amended by the Legislature many times on an ad hoc basis but they have not been systematically assessed since they were enacted nearly 25 years ago. The underlying assumptions built into them have never been tested. Their relationship to current knowledge about the effectiveness of criminal justice practices and to evolving public attitudes toward crime and punishment have not been considered.

The research reported here provides a window into what a commission might learn. It provides suggestions about issues a commission might consider. But it is no substitute for a cross-professional commission with a broad mandate, full-time professional staff, and adequate resources. Such a commission could ultimately make sentencing in Michigan fairer, more consistent, and more cost-effective. That is, it could utilize all we have learned over the past 40 years and make the vision of the Michigan Felony Sentencing Project a reality.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Endnotes

The Point of This Report

[1] See discussions in Chapter One of the Michigan Sentencing Guidelines Commission (1995–2002) and the Michigan Criminal Justice Policy Commission (2015–2019).

[2] Michigan's sentencing scheme is called "indeterminate" because the minimum sentence establishes only when the person *can* be released, not when they will be.

- The minimum sentence determines when the parole board obtains jurisdiction to release a prisoner. A defendant may serve more time than their minimum requires but cannot serve less without a commutation of sentence by the governor.
- Depending on the amount of time between the minimum and maximum sentences—what prisoners call the "tail" of the sentence—there may be years or decades within which the board can decide whether to grant release.
- Historically, prisoners around the country had their minimum sentences reduced by "time off for good behavior" or "good time." In most jurisdictions, some amount of good time—ranging from very minimal to very generous—is still available. However, Michigan wholly eliminated what were then known as disciplinary credits after the adoption of "truth in sentencing" in 1998. Prisoners sentenced since then must serve every day of their minimum sentences.
    - With credits, on average prisoners not sentenced to life terms had served 88% of their minimums.
    - In 2019 the average minimum sentence was 10.8 years. Disciplinary credits would have reduced the average time to parole eligibility to 9.3 years. Thus the elimination of credits increased the average length of stay by 1.3 years.

[3] The other major exception is the small but important group of offenses with mandatory sentences, particularly life without parole for first-degree murder and the consecutive two-year term required for convictions of "felony firearm" (possessing a firearm during the commission of a felony). In theory at least, indeterminate sentencing defines appropriate punishment in light of the defendant's background, not just the nature of the criminal behavior. With mandatory terms, judges have no discretion to consider individual circumstances and tailor the sentence accordingly. In 2019, 3,899 people were serving mandatory life without parole.

[4] Beggin, Riley. 2019. "Michigan Jails Fill as Crime Sinks and Nobody Seems to Know Why." https://www.bridgemi.com/michigan-government/michigan-jails-fill-crime-sinks-and-nobody-seems-know-why

[5] Risko, Robin. 2021. *Budget Briefing: Corrections.* Michigan House Fiscal Agency. https://www.house.mi.gov/hfa/PDF/Briefings/Corrections_BudgetBriefing_fy20-21.pdf

[6] These have included such diverse matters as asset forfeiture, cash bail, parole criteria, expungement, and the age of criminal responsibility.

Section I. History and Context

Chapter One. The Evolution of Sentencing Guidelines

[1] Ballotpedia. N.d. "Michigan Indeterminate Sentence Proposal, 1902. Retrieved Mar 2, 2021. https://ballotpedia.org/Michigan_Indeterminate_Sentence,_Proposal_2_(1902)

[2] See summary of case law in *People v. Coles*, 417 Mich 523 (1983).

[3] *People v. Tanner,* 387 Mich 683, 689 (1972).

[4] It later held that the rule does not apply to life-maximum offenses. *People v. Powe,* 469 Mich 1032 (2004).

[5] Zalman, Marvin, Charles W. Ostrom, Phillip Guilliams, and Garret Peaslee. 1979. *Sentencing in Michigan: Report of the Michigan Felony Sentencing Project*. NCJRS. https://www.ojp.gov/pdffiles1/Digitization/60907NCJRS.pdf

[6] *Ibid* at 174.

[7] *Ibid* at 265–266.

[8] Their findings will be discussed in more detail in Chapter Three.

[9] Zalman supra note 5 at 295.

RECEIVED by MSC 4/12/2022 1:07:49 PM

[10] A subsequent legislative analysis noted: "The [judicial sentencing] guidelines were designed to reduce disparity in sentencing from county to county and region to region by mirroring the existing sentencing practices of judges across the state at the time the guidelines were implemented. They were developed using the results of research on sentencing patterns of judges throughout Michigan, and attempt to capture the typical sentence for similar types of offenses and offenders." House Legislative Analysis, SB 826, HB 5419, and HB 5398 (Revised Second Analysis), September 23, 1998, 2.

[11] McComb, James A. 1988. "An Overview of the Second Edition of the Michigan Sentencing Guidelines." *Michigan Bar Journal*, 67:863–867 at 863.

[12] See Michigan Judicial Institute [MJI]. 1983–1988. "Michigan Sentencing Guidelines Manual." Retrieved on Mar 2. 2021. https://mjieducation.mi.gov/documents/felony-sentencing-resources/59-1983-1988-sgm/file The groups were assault, burglary, criminal sexual conduct, drug, homicide, larceny, robbery, fraud, property destruction, weapons possession, and negligent homicide.

[13] Deming, Sheila Robertson. 2000. "Michigan's Sentencing Guidelines." *Michigan Bar Journal*, 79(6):652-655 at 652.

[14] *People v. Ewing*, 435 Mich 443 (1990).

[15] The directions in the 1988 Sentencing Guidelines Manual described the process as follows:

> "6. Do not score any prior felony convictions, misdemeanors or juvenile delinquency adjudications that precede conviction-free periods of 10 years or more. A conviction-free period exists if more than 10 years have elapsed between the discharge date from any conviction or adjudication and the commission of the next offense that results in a conviction… To determine whether the decay factor is applicable complete the following steps.
>
> a. Start from the date of the instant offense and go back in time to the most recent prior conviction and determine the discharge date. If there is less than 10 years between the two dates, the most recent prior conviction is to be scored. If the gap is more than 10 years, that prior conviction and any previous convictions are not to be scored.
>
> b. If there is a prior conviction within the most recent 10 year period, start from the offense date of the most recent prior conviction and go back in time to the discharge date of the next prior conviction. If there is less than 10 years between the two dates, the next prior conviction is to be scored. If there is more than a 10 year gap, that prior conviction and all previous convictions are not to be scored.
>
> c. Continue in a similar fashion until there is a 10 year gap or until there are no further prior convictions."

Michigan Judicial Institute. 1988. "Michigan Sentencing Guidelines (Second Edition)." pp. 6–7. Retrieved on March 2, 2021. https://mjieducation.mi.gov/documents/felony-sentencing-resources/60-1988-2nded-sgm/file

[16] MJI supra note 12 at 154.

[17] McComb supra note 11 at 864.

[18] *Ibid*. The revised guidelines also reduced the number of crime groups from 11 to 10 by dropping negligent homicide as a separate group.

[19] *Ibid* at 866.

[20] *Ibid* at 867.

[21] Ostrom, Charles W., Brian J. Ostrom, and Matthew Kleiman. 2004. "Judges and Discrimination: Assessing the Theory and Practice of Criminal Sentencing." NCJRS, at 162.https://www.ojp.gov/pdffiles1/nij/grants/204024.pdf

[22] 417 Mich 523 (1983).

[23] *People v. Milbourn*, 435 Mich 630 (1990).

[24] *Ibid* at 653.

[25] *Ibid* at 659.

[26] *Ibid.*

[27] A third judge, Paul Maloney, was named as a public member and appointed chair.

[28] Index crimes include murder, rape, robbery, aggravated assault, burglary, larceny, motor vehicle theft, and arson.

[29] The Disaster Center. N. d. "Michigan Population and Rate of Crime per 100,000 People, 1960-2019." Retrieved on Mar 2, 2021. http://www.disastercenter.com/crime/micrime.htm

[30] Intermediate sanctions are required if the high end of the guidelines range is 18 months or less.

[31] 498 Mich 358 (2015).

RECEIVED by MSC 4/12/2022 1:07:49 PM

[32] These definitions were provided by the Supreme Court in *People v. Babcock*, 469 Mich 247 (2003). The *Babcock* Court stressed that the principle of proportionality remained relevant as the standard against which the decision to depart was to be assessed. In establishing the sentencing guidelines, the Legislature was subscribing to the "deeply rooted" belief that the punishment should fit both the crime and the criminal. The scoring of offense and prior record variables reflects this belief. Therefore, a departure must contribute to making a sentence more proportionate than a sentence within the guidelines range.

Substantial and compelling reasons must justify not only the fact of *a* departure, but also the extent of the particular departure imposed. As the Supreme Court explained in *People v. Smith*, 482 Mich 292 (2008), a departure that is too extreme will violate the principle of proportionality.

[33] *People v. Hendrick*, 472 Mich 555 (2005).

[34] Straddle cells are those in which the high end of the guidelines range exceeds 18 months but the low end is 12 months or less.

[35] The Council of State Governments found that prison sentences imposed in 2012 were distributed among the crimes classes as follows:

- M2 (life or any term) 2%
- Class A (life or any term) 11%
- Class B (20 years) 11%
- Class C (15 years) 14%
- Class D (10 years) 16%
- Class E (5 years) 27%
- Class F (4 years) 7%
- Class G (2 years) 10%
- Class H (intermed. sanc.) 1%

Council of State Governments Justice Center [CSG]. 2014. "Report Technical Appendix: Compilation of Michigan Sentencing and Justice Reinvestment Analyses." p. 29.
https://council.legislature.mi.gov/Content/Files/mlrc/MichiganReportTechnicalAppendix.pdf

[36] 1st judicial guidelines, assault with intent to murder, high offense, low priors.

[37] Yantus, Anne. 2014. "Sentence Creep: Increasing Penalties in Michigan and the Need for Sentencing Reform." *University of Michigan Journal of Law Reform*, 47(3):645–696 at 685.

[38] MCL 769.10.

[39] MCL 769.11.

[40] MCL 769.12.

[41] Deming supra note 13.

[42] Act of Mar. 7, 2002, No 31, 2002 Mich. Pub. Acts 68, 74 (repealing Mich. Comp. Laws Secs. 769.32-33)

[43] Yantus supra note 37 at 685, nn 273.

[44] *Ibid* at 652–658.

[45] *Ibid* at 664–665.

[46] *Ibid* at 666–667.

[47] MCL 768.7a. Mandatory consecutive sentences if defendant was inmate of or on escape from a penal institution or was on parole when they committed the offense. See also MCL 768.7b regarding offenses committed while a felony charge is pending. Consecutive sentencing is mandatory if new crime is a major controlled substance offense; otherwise consecutive sentencing is discretionary.

[48] MCL 750.227b.

[49] MCL 333.7401(3).

[50] Yantus supra note 37 at 681–683.

[51] *People v. Miles*, 454 Mich 90 (1997).

[52] *People v. Tommy Brown*, Michigan Supreme Court No. 147759 (order remanding, 3/26/14) (approving *People v. Ryan*, 295 Mich App 388 (2012)). Brown was also convicted of four other counts of first-degree CSC for other incidents involving the same teenaged victim. The 40–60 year terms imposed for those counts had to run concurrently. It appears from the Michigan Department of Corrections' Offender Tracking Information System (OTIS) that when Brown was resentenced on August 14, 2014, he received seven terms of 25–50 years to be served concurrently since his earliest release date is now January 4, 2037.

RECEIVED by MSC 4/12/2022 1:07:49 PM

[53] *People v. Christopher Lynd Barron,* unpublished opinion of Court of Appeals issued July 2, 2019 (Docket No. 339508); Christenson, Trace. 2017. "Bellevue Man Faces Up to 160 Years in Prison," *Lansing State Journal*, July 15, pp. 5A.

[54] Yantus supra note 37 at 683, nn 264.

[55] The Court of Appeals acknowledged the illogic in *People v. Ryan,* supra note 52.

[56] 317 Mich App 649 (2016).

[57] On remand, the trial court made two of the sentences consecutive and ordered the rest to run concurrently to each other and the first two, reducing the aggregate minimum sentences to 22 years. The justification given included the defendant's extensive and violent criminal history, his failure to be gainfully employed, his manipulation of addicts and his 18-year old girlfriend to sell heroin, and his long history of dealing. The Court of Appeals affirmed. *People v. Norfleet (On Remand),* 321 Mich. App. 68 (2017).

[58] A mandatory minimum assumes an indeterminate sentence with both minimum and maximum terms. This is distinct from a "flat" or determinate sentence for which the Legislature has mandated a single term of years. While some states use determinate sentences routinely, in Michigan, the primary instances are the consecutive two-year term for possessing a firearm during the commission of a felony, MCL 750.227b and life without parole for first-degree murder, MCL 750.316.

[59] Yantus supra note 37 at 678.

[60] Life without parole is now also mandatory for a defendant convicted of first-degree CSC against a child under age 13, if the defendant has a prior conviction of first, second or third-degree CSC involving a child. MCL 750.520b(2)(c).

[61] MCL 769.12.

[62] Yantus supra note 37 at 680.

[63] 2014 PA 465.

Chapter Two. Changing Views of How Long is Long Enough

[1] For more details on the distribution of minimum sentences for these offenses in the decades from 1970-2009, see Citizens Alliance on Prisons and Public Spending. Fall 2012. "Length of Stay Drives Prison Costs." *Consensus* at p. 5.

[2] Aebi, Marcelo F., and Mélanie M. Tiago. 2018. "Prisons and Prisoners in Europe 2018: Key Findings of the SPACE I Report." *Council of Europe.* 1–17 at p 4. http://wp.unil.ch/space/files/2019/06/Key-Findings_190611-1.pdf Norway and many other countries allow for sentences to be extended or converted to "preventive detention" if the person is determined to be a continuing threat to public safety.

[3] Mauer, Marc. 2017. "Incarceration Rates in an International Perspective*." Oxford Research Encyclopedia of Criminology and Criminal Justice*, 1–20. https://oxfordre.com/criminology/view/10.1093/acrefore /9780190264079.001.0001/acrefore-9780190264079-e- 233#acrefore-9780190264079-e-233-bibItem-0036.

[4] Index crimes include murder, rape, robbery, aggravated assault, burglary, larceny, motor vehicle theft and arson.

[5] Comparative averages for first-degree CSC offenses cannot be calculated because they are displayed differently in the MDOC reports for the years in question.

The Council of State Governments (CSG) made similar findings when it examined Michigan felony sentences imposed during the five-year period from 2008–2012. CSG found that increases were most marked for cases on the second-degree murder and Class A grids, (i.e. the life-maximum offenses).

The average minimum sentence for murder grew from 277.9 months to 309.6 months, an increase of 11.4%, even though the average offense variable (OV) score grew only marginally and the average prior record variable (PRV) score actually declined.

The average minimum sentence for A grid offenses grew from 121.4 months to 132.7 months, a 9.3% increase, although the average OV score had not changed and the average PRV score declined slightly. See Council of State Governments, supra note 35 in ch 1 at 25, 27.

[6] Leigh, Courtney, Sarah Eppler-Epstein, Elizabeth Pelletier, Ryan King, and Serena Lei. 2017. "A Matter of Time: The Causes and Consequence of Rising Time Served in America's Prisons." Urban Institute: Justice Policy Center, 1–4. https://apps.urban.org/features/long-prison-terms/trends.html

[7] Gottschalk, Marie. 2016*. Caught: The Prison State and the Lockdown of American Politics*. Princeton: Princeton University Press.

[8] Prescott, James J., Benjamin Pyle, and Sonja B. Starr. 2019. "Understanding Violent-Crime Recidivism." *Notre Dame Law Review*, 95(4):1643-1698 at 1697.

RECEIVED by MSC 4/12/2022 1:07:49 PM

[9] Mauer, Marc, Ashley Nellis, and Kerry Myers. 2018. *The Meaning of Life: The Case for Abolishing Life Sentences*. New York: The New Press.

[10] Pfaff, John. 2017. *Locked In: The True Causes of Mass Incarceration and How to Achieve Real Reform.* New York: Basic Books.

[11] King, Ryan, Bryce Peterson, Brian Elderbroom, and Elizabeth Pelletier. 2015. "Reducing Mass Incarceration Requires Far-Reaching Reforms." Retrieved Mar 2, 2021. https://apps.urban.org/features/reducing-mass- incarceration. A 15% reduction in the length of stay of those convicted of all nonviolent offenses would reduce the baseline population by 4.5%.

[12] Austin, James, Vincent Schiraldi, Bruce Western, and Anamiki Dwivedi. 2019. "Reconsidering the "Violent Offender,"" Retrieved Mar 2, 2021. https://squareonejustice.org/wp-content/uploads/2019/09/executive-session-pdf-Reconsidering-the-violent-offender-report-ONLINE_FINAL.pdf

[13] Michigan Department of Corrections. 1993. "1991 Statistical Report." at Table B2c:Commitments and Minimum Term Distribution for Men and Women-Assaultive Offenses.

[14] Michigan Department of Corrections. 2021. "2019 Statistical Report." at Table B2c: 2019 Commitments by Minimum Term Distribution-Assaultive Offenses. https://www.michigan.gov/documents/corrections/MDOC_2019_Statistical_Report_717026_7.pdf

[15] Risko supra note 5 in The Point of This Report.

[16] *Ibid* at 31.

[17] *Ibid* at 32.

[18] Legislation to permit the release of medically frail prisoners to approved nursing homes was amended to exclude anyone serving for first-degree murder or first-degree CSC. Since these people make up the vast majority of the elderly prisoner population, very few of the medically frail will actually be eligible for medical paroles. MCL 791.235(10).

[19] Rhodes, William, Gerald Gaes, Jeremy Luallen, Ryan Kling, Tom Rich, and Michael Shively. 2016. "Following Incarceration, Most Released Offenders Never Return to Prison." *Crime & Delinquency*, 62(8):1003–1025. The authors do not identify the 17 states or the number of people in the dataset. The study does not distinguish between returns to prison for new crimes and returns for parole violations. The authors emphasize that their offender-based method of counting produces a lower recidivism rate than studies that count events, i.e. each return to prison, because the latter method counts people who return multiple times each time they appear in the data.

[20] Levine, Barbara R. 2009. *Denying Parole at First Eligibility: How much public safety does it actually buy?* Retrieved Mar 2, 2021. https://www.safeandjustmi.org/wp-content/uploads/2009/08/Denying_parole_at_first_eligibility.pdf

[21] Michigan Department of Corrections. 2021. "2019 Statistical Report." At table D3: Three-Year Follow-Up Outcomes of Offenders Who Paroled in 2001 to 2016 by Year.

[22] For a review of the research see Ulmer, Jeffrey T., and Darrell J. Steffensmeier. 2014. "The Age and Crime Relationship: Social Variation, Social Explanations." Pp. 377-396 in *The Nurture Versus Biosocial Debate in Criminology: On the Origins of Criminal Behavior and Criminality*. Thousand Oaks: SAGE Publications Inc.

[23] Levine supra note 20.

[24] *Ibid* at 23–27.

[25] Levine, Barbara and Elsie Kettunen. 2014. "Paroling People Who Committed Serious Crimes: What is the Actual Risk?" Retrieved Mar 2, 2021. https://www.prisonpolicy.org/scans/cappsmi/CAPPS_Paroling_people_who_committed_serious_crimes_11_23_14.pdf

[26] Justice Policy Institute. 2018. "The Ungers, 5 Years and Counting: A Case Study in Safely Reducing Long Prison Terms and Saving Taxpayer Dollars." Retrieved Mar 2, 2021. http://www.justicepolicy.org/uploads/justicepolicy/documents/The_Ungers_5_Years_and_Counting.pdf

[27] Prescott supra note 8.

[28] *Ibid* at 1687 [Table 5]. Re-incarceration rates were 4.9% for the rape/sex assault group, 9.9 for the assault group, and 11.0% for the robbery group.

[29] Levine supra note 20 at 43–48. See also Rhodes, William, Gerald G. Gaes, Ryan Kling, and Christopher Cutler. 2017. "The Relationship Between Prison Length of Stay and Recidivism: A Study Using Regression Discontinuity with Multiple Break Points." *Criminology & Public Policy*, 17(3):731–769. (Statistical analysis of federal data finds very little impact of length of stay on recidivism).

RECEIVED by MSC 4/12/2022 1:07:49 PM

Chapter Three. Prior Research on Michigan Sentencing Guidelines

[1] Zalman supra note 5 in ch 1 at 72.

[2] *Ibid* at 169. Offender variables included not only prior record and relationship to the criminal justice system but also a wide variety of information judges may find relevant, such as residential stability, employment, education, mental health, substance abuse, and "good moves since arrest." (This sort of background information that is found in confidential presentence reports but not scored in the legislative guidelines was not available for the current research.) Age, gender, and race were analyzed separately.

[3] *Ibid* at 191–92.

[4] *Ibid* at 186–87.

[5] *Ibid* at 213.

[6] *Ibid* at 261.

[7] *Ibid* at 263.

[8] *Ibid* at 234.

[9] *Ibid* at 234. For instance, a non-white using a weapon to commit a sex offense received, on average, a minimum of 105 months where a white received 42 months. If the victim of a sex offense was a stranger, the non-white's sentence would increase by 54 months but the white's sentence would decrease by four months. *Ibid* at 250-252.

[10] *Ibid* at 269.

[11] Ekpunobi, Abel E 1999. "Judicial Decision Making Under Michigan Sentencing Guidelines." *Dissertations*, 1504.

[12] The actual identities of the counties were confirmed by Dr. Ekpunobi in an e-mail dated Feb. 3, 2021.

[13] The study also examined the race and gender of the sentencing judges, with mixed results. For further discussion of the 2nd judicial guidelines, see the analysis of the 1995 data used by the Michigan Sentencing Commission in Ostrum supra note 22 in ch 1.

[14] Ostrom, Brian J., Charles W. Ostrom, Roger A. Hanson, and Matthew Kleiman. 2008. Assessing Consistency and Fairness in Sentencing: A Comparative Study in Three States. *NCJRS*. https://www.ojp.gov/pdffiles1/nij/grants/223854.pdf

A short summary of the findings was published under the same title in Aug. 2008. References here are to the full report.

[15] *Ibid* at 82.

[16] *Ibid* at 80.

[17] *Ibid* at 85. Other than second-degree murder, which has its own grid, NCSC researchers did not distinguish among specific offenses in its analyses. Because they had only 77 murder cases in their dataset, for some analyses they collapsed the cases on the M2 and A grids.

[18] *Ibid* at 118–19.

[19] *Ibid* at 89.

[20] *Ibid* at 170–71.

[21] *Ibid* at 171.

[22] *Ibid* at 187.

[23] *Ibid* at 174.

[24] *Ibid* at 177.

[25] *Ibid* at 177.

[26] *Ibid* at 185.

[27] *Ibid* at 186.

[28] *Ibid* at 188.

[29] *Ibid* at 184–85.

[30] *Ibid* at 188.

[31] *Ibid* at 210.

[32] *Ibid* at 206.

[33] *Ibid* at 195.

[34] *Ibid* at 188.

[35] *Ibid* at 201.

[36] *Ibid* at 210.

[37] *Ibid* at 211.

[38] *Ibid* at 193.

RECEIVED by MSC 4/12/2022 1:07:49 PM

[39] *Ibid* at 300.
[40] *Ibid* at 301.
[41] *Ibid* at 302–03.
[42] *Ibid* at 303.
[43] CSG supra note 35 in ch 1 at 16–17.
[44] *Ibid* at 23.
[45] *Ibid* at 24.
[46] *Ibid* at 17.
[47] *Ibid* at 30.
[48] Criminal Justice Policy Commission [CJPC]. 2019. *Evaluation of Straddle Cell Sentencing in Michigan: Summary Report and Recommendations*. http://council.legislature.mi.gov/Content/Files/cjpc/CJPCStraddleCellSentencingLegislators.pdf

Section II. The Current Research

Chapter Five. How Sentences Grew Longer and More Numerous

[1] For detailed discussions of judicial views, parole policy changes, and lifer release patterns, see, respectively:

- Prisons and Corrections Section, State Bar of Michigan. 2002. "What Should 'Parolable Life' Mean? Judges Respond to the Controversy." https://www.safeandjustmi.org/wp-content/uploads/2002/03/What_should_parolable_life_mean.pdf
- Levine, Barbara.. 2014. "Parolable Lifers in Michigan: Paying the Price of Unchecked Discretion." https://www.prisonpolicy.org/scans/cappsmi/Parolable-Lifers-in-Michigan-Paying-the-price-of-unchecked-discretion.pdf
- Citizens Alliance on Prisons and Public Spending. 2006. "When 'Life' Did Not Mean Life: A Historical Analysis of Life Sentences Imposed in Michigan since 1900." https://static.prisonpolicy.org/scans/cappsmi/When%20life%20did%20not%20mean%20life%20for%20web.pdf

[2] Michigan Department of Corrections. 2020. "Report to the Legislature, Pursuant to Article 2, Public Act 166 of 2020, Section 904." https://www.michigan.gov/documents/corrections/Sec_904_713467_7.pdf

[3] The number of sentences in each offense group by habitual offender status under the legislative guidelines was:

	Non-Habitual	2nd Habitual	3rd Habitual	4th Habitual
M2	1,723	108	70	78
AWIM	1,398	123	99	126
CSC	1,976	190	100	164
RA	8,883	643	463	807
Total	13,980	1,064	732	1,175

Chapter Six. How the Guidelines Promote Longer Sentences

[1] The groups are crimes: against a person, against property, involving a controlled substance, against public order, against public safety and against public trust.

[2] The CSG observed that Michigan's guidelines are based on more offense groupings and incorporate more offense and prior record variables than other guidelines states, leading to far more cells. Reynolds, Carl, Andy Barbee, Ellen Whelan-Wuest, and Cassandra Warney. "Sentencing and Justice Reinvestment Initiative." Presented to Michigan Law Revision Commission, March 19, 2014, at p. 2. https://csgjusticecenter.org/wp-content/uploads/2020/10/MLRC-Pres-May-13-2014.pdf .Yet the overlap in allowable sentence lengths means the "complex scoring yields illusory precision." *Ibid* at 12.

The Robina Institute has also noted that Michigan's number of grids is unique. Frase, Richard S., Julian, V. Roberts, Rhys Hester, and Kelly Lyn Mitchell. 2015. Criminal History Enhancements Sourcebook. *Robina Institute of Criminal Law and Criminal Justice*. https://robinainstitute.umn.edu/sites/robinainstitute.umn.edu/files/criminal_history_enhancement_web2_0.pdf. Most

RECEIVED by MSC 4/12/2022 1:07:49 PM

jurisdictions that use guidelines grids have one or two. Frase at 10. The authors of a Robina analysis that attempted to compare the impact of criminal history scores in guidelines jurisdictions had to exclude Michigan "due to complications related to that state's use of nine separate grids. Frase at 21.

[3] *Ibid.*

[4] *Ibid* at 11–14.

[5] *Ibid* at 97–104.

[6] *Ibid* at 20–21 and 25–27.

[7] Table A2, 2019 Criminal Court Dispositions by Offense and Type of Disposition — All Offenses

[8] MCL 333.7403(2)(b)(i): 10-year maximum, 3,591 dispositions, 14.1% to prison.

[9] MCL 333.7401(2)(a)(iv): 20 year maximum, 1,992 dispositions, 31.4% to prison.

[10] MCL 750.84: 10-year maximum, 1,064 dispositions, 47.7% to prison.

[11] MCL 750.110: 10-year maximum, 1,022 dispositions, 31.8% to prison.

[12] MCL 750.249: 14-year maximum, 1,016 dispositions, 23.7% to prison.

[13] Consider a scenario that could occur under MCL 750.380, which raises the maximum sentence for malicious destruction of a house, barn, or other building from five to 10 years if the defendant has two prior convictions under the same provision. Assume that Charlie has no prior record of any kind but has had a very hard time dealing with a divorce and the fact that his wife got the farm they built up together. On three separate occasions, he was so angry that he damaged property worth a little over $1,000. The first time he was placed on probation. The second time he was placed on probation again but had to serve 30 days in jail. The third time, the maximum became 10 years. And this time he also was convicted of felonious assault for hitting his ex-wife's new boyfriend who tried to intervene. Charlie's sentence will be on the D grid because of the 10-year maximum. He will receive 10 points under PRV 2 for the two prior low severity felonies, 10 points under PRV 6 because he was on probation, and 10 points under PRV 7 because of the concurrent assault conviction. This will place him at PRV Level D. He will receive 10 points under OV 13 because he now has a pattern of three or more property crimes and five points under OV 16 because the property was valued over $1,000. This puts him at OV Level II. The D-II cell has a recommended minimum prison sentence of 5–23 months. Charlie could also have his sentenced enhanced as a fourth habitual offender.

[14] Frase supra note 2 at 48.

[15] *Ibid* at 84.

[16] MCL 800.283a.

[17] It should be noted that a 2004 change in Michigan law defining when multiple convictions are barred by the state constitutional protection against double jeopardy allowed for more concurrent convictions to arise from a single transaction. This, in turn, increased the number of convictions available for scoring under PRV 7. *See People v. Nutt,* 469 Mich 565 (2004).

[18] For example, the requisite 50 points can be reached with two prior low-level felonies (10), three low severity juvenile adjudications (10), three misdemeanors (10), being on probation for one of the felonies (10), and one felony conviction that is concurrent with the sentencing offense.

[19] Zalman supra note 5 in ch 1.

[20] *Guidelines Manual*, online version current through 1/20/21 Step II. B. https://mjieducation.mi.gov/documents/sgm-files/94-sgm/file

[21] The only exception is that five points are not to be scored for displaying or implying a weapon if the conviction offense was armed robbery or felonious assault.

[22] The American Law Institute. 2017. *Model Penal Code: Sentencing Proposed Final Draft.* Sec. 7.03 (3) *Eligible Sentencing Considerations.* https://robinainstitute.umn.edu/sites/robinainstitute.umn.edu/files/mpcs_proposed_final_draft.pdf

[23] *Ibid.*

[24] *People v. Hardy, People v. Glenn,* 494 Mich 430, 442 (2013).

[25] *Ibid.* The Legislature amended OV 7 effective Jan. 2016 to supersede *Hardy* in part. It added the phrase "similarly egregious" before "conduct designed to substantially increase the victim's fear." That the new language failed to provide a satisfactory solution to the problem of excessively scoring 50 points is evident in *People v. Lydic*, __Mich App__ (COA No. 349216, rel'd 1/28/21). During a domestic dispute, the defendant strangled his girlfriend with a belt and said that her young son would come home to find her dead. The defendant received a below guidelines sentence of 17 months to 10 years. The trial court felt compelled to score 50 points under OV 7 but then chose to depart downward in part because it believed that score was too harsh under the circumstances. The Court of Appeals upheld both the scoring and the departure.

[26] *People v. Barrera,* 500 Mich 14 (2017).

RECEIVED by MSC 4/12/2022 1:07:49 PM

[27] Ten points are also scored if four to 19 victims "were placed in danger of property loss"; 25 points are scored if 20 or more victims "were placed in danger of property loss." This has particular impact if the offense is breaking and entering or home invasion, when multiple victims not present during the crime might have lost property.

[28] Model Penal Code, supra note 22 at Sec. 7.03 (2)(b). See also Reitz, Kevin. 1993. "Sentencing Facts: Travesties of Real-Offense Sentencing." *Stanford Law Review*, 45(3):523-573. https://doi.org/10.2307/1229007

[29] *People v. Underwood*, 278 Mich App 334, 339-340 (2008) (absent an express statutory exclusion, 10 points may be awarded for perjury under OV 19 although interference with the administration of justice was inherent in the perjury conviction for which the defendant was being sentenced).

[30] *People v. Dickinson*, 321 Mich App 1, 24 (2017) (25 points properly awarded under OV19 when conviction offense was furnishing a controlled substance to a prisoner).

[31] *People v. Hamin Lorenzo Dixon,* __Mich App __ (COA No. 349631, 9/10/20) (order directing argument on app for lv., SC. No. 16221, 4/9/21) (25 points properly awarded under OV 19 where defendant pled guilty to attempted possession of a cell phone by a prisoner).

[32] The frequencies in this section include both non-habitual and habitual sentences. If they have the same guidelines scores, habitual and non-habitual sentences fall in the same cells; it is only the maximum-minimum that changes depending on habitual status.

[33] MCL 769.33(1)(e)(vii).

[34] Ostrom, Charles W. 1997. *Technical Report, Pilot Project: Using the MSP Draft Guidelines in Six Counties – Genesee, Kalamazoo, Kent, Macomb, Marquette, Oakland – and Recorder's Court*.

[35] *Ibid*; Maloney, Paul L., Hilda Gage, Mark Murray, and Carlo Ginotti. 1997. *Report of the Michigan Sentencing Guidelines Commission*. Appendix C.

[36] *Ibid* at 7.

[37] *Ibid* at 8.

[38] *Ibid* at 29.

[39] *Ibid* at 31.

[40] *Ibid* at 28.

[41] 482 Mich 41 (2008).

[42] For more detail about the regression analyses, see Appendix B.

[43] See Appendix B for details about how changing the order the variables are entered into the model can change the amount of variance explained.

Chapter Seven. How the Guidelines Promote Disparity

[1] Maloney, Paul L. 1999. "The Michigan Sentencing Guidelines," *Thomas M. Cooley Law Review*. 16(1):13-26. https://heinonline.org/HOL/LandingPage?handle=hein.journals/tmclr16&div=8&id=&page=

[2] *Ibid* at 21.

[3] *Ibid.*

[4] Michigan Supreme Court Guidelines Program. 1994. "Departure Profile for Mandatory Guideline Period. (01/01/94-12/31/94)."

[5] Maloney supra note 35 in ch 6 at 15.

[6] Ostrom supra note 14 in ch 3 at 84. Looking at sentences imposed from 2008-2012, the Council of State Governments (CSG) concluded that the ranges for prison sentences are wide and that the sentences imposed are spread across the ranges and beyond. It used as an example one cell on the D grid for the offense of delivering drugs < 50 grams. The low end of the cell was 10 months and the high end was 23 months. It found that only 58% of the sentences were within the range. CSG supra note 36 in ch 1 at 23.

[7] Michigan Supreme Court Guidelines Program, supra note 4.

[8] The Sentencing Commission had projected that the probable average sentence on the M2 grid would be at the cell midpoint and the probable average sentence on the A grid would be midway between the cell minimum and the cell midpoint. Maloney supra note 35 in ch 6 at 15.

Chapter Eight. How the Actual Minimum Sentences Relate to the Recommended Ranges

[1] 2006 PA 169, effec. Aug. 28, 2006.

[2] Significant at .001 level.

RECEIVED by MSC 4/12/2022 1:07:49 PM

[3] Another small but noticeable phenomenon is that in the A-I cell on three offense grids, the median sentence actually exceeded the high end of the recommended range, i.e., the median was an upward departure.

- On the M2 grid, where the max-min of the A-I cell is 150 months, the median for the 14 sentences in that cell was 168 months.
- For the offenses on the A grid, where the high end of the A-I range is 35 months, the median sentences were 180 months for assault and 60 months for CSC.

These medians resulted from upward departures in 28 individual sentences across the three offense groups. Our review suggests that in the eight murder cases, judges may simply have felt that 12.5 years was not enough punishment for the crime. In the assault and CSC cases, it appears that a frequent explanation for why judges saw a top of the guidelines sentence as insufficient was the existence of other sentences. This might be multiple sentences of the same type or a sentence for another offense for which the defendant was receiving a longer sentence, as when someone with no prior record was convicted of multiple counts of CSC or assault and murder out of the same incident.

[4] The principle exception is the robbery grid. Because so many medians are at the lowest end of their respective ranges, the starting points and medians of redundant cells along each diagonal are frequently the same.

Chapter Nine. Systemic Factors: Method of Conviction

[1] 397 U.S. 742 (1970). Brady pled guilty to avoid imposition of the death penalty, which was allowed under the applicable statute only if recommended by a jury. Nearly a decade later, the Supreme Court held in another case that this penalty provision unconstitutionally burdened the right to a jury trial because it required the defendant to risk death in order to exercise the right. However, the Court upheld Brady's guilty plea on the rationale that it was voluntary when made and was not invalidated by a subsequent change in the law.

[2] For more detailed discussion, see Grossman, Steven P. 2005. "An Honest Approach to Plea Bargaining." *American Journal of Trial Advocacy*, 29: 101-136. See also Smith, Justin M. 2020. "Disparity in Context: Judges' Perspectives on Disparities in a Sentencing Guideline System." *Journal of Qualitative Criminal Justice & Criminology*, 8(2):1-36 10.21428/88de04a1.52cab578 at pp. 21–23 for interviews with Michigan judges about disparities caused by plea bargaining and impact of witnesses at trial.

[3] Grossman supra note 2 at 130, nns omitted. See also *People v. Byrd*, 12 Mich App 186, 194 (1968) (Levin, J. concurring). Plea bargaining only works if defendants know that sentences will be heavier after trial. In Michigan in 1982, the Supreme Court decided the companion cases of *People v. Biggs* and *People v. Killebrew*, 416 Mich 189 (1982). The Court held that while the practice of charge bargaining with the prosecutor, which is intended to constrain the length of the potential sentence, had long been accepted, the practice of bargaining directly for imposition of a particular sentence or for a recommendation from the prosecutor that the judge impose a particular sentence, had not been expressly recognized as appropriate. The Court prohibited judicial initiation of or participation in sentence negotiations. However, it adopted a process by which judges can approve or reject sentence agreements or recommendations negotiated by the parties so long as the defendant has the opportunity to withdraw his plea if the bargain is rejected.

Ten years later, the Court extended this holding in *People v. Cobbs*, 443 Mich 276 (1993) to allow judges, upon request of a party, to state on the record what the sentence would be based on the judge's knowledge of the case at that point. The judge's preliminary evaluation does not bind the court to impose that sentence, but if the defendant pleads guilty in reliance on an estimate that the judge ultimately decides to exceed, the defendant is entitled to withdraw the plea.

[4] *Biggs* supra note 3. The trial judge rejected the sentence recommendation and imposed a sentence of 20–40 years. Bargains may also be struck about how guidelines variables will be scored or even which grid will be used, e.g., whether a second-degree murder will be scored as manslaughter on the C grid.

[5] The decision to go to trial requires a calculation of the likelihood of conviction. A lawyer may recommend acceptance of a bargain because conviction appears likely, regardless of how strongly the defendant asserts innocence. However, there is also an emotional component to the decision. Defendants, especially young ones, may have an unrealistically optimistic view of the trial process or an inability to accept the consequences of the offered plea. Defendants may also reject bargains they feel are unfair, either because they believe the punishment is too harsh for the crime or because they see it as harsher than normally imposed in similar circumstances.

The ultimate example of such a decision is *Bordenkircher v. Hayes*, 434 U.S. 357 (1978). The defendant was charged with uttering a forged instrument in the amount of $88.30 which, under then-applicable Kentucky law, was punishable by 2–10 years in prison. At a prehearing conference the prosecutor offered to recommend a sentence of five years if Hayes would plead but also threatened to seek indictment as a habitual criminal if Hayes refused to

RECEIVED by MSC 4/12/2022 1:07:49 PM

"save the court the inconvenience and necessity of a trial." With two prior convictions, Hayes faced a mandatory sentence of parolable life. He refused the plea offer, was found guilty on the principle charge by a jury, and received the life term.

The U.S. Supreme Court upheld the result as a product of the "give and take" of plea bargaining and observed that confronting a defendant with the risk of more severe punishment was an inevitable attribute of a system that encourages negotiated pleas. So long as the defendant was informed of the risks when he made his decision to plead not guilty, the threat of increased punishment for exercising the right to trial did not violate due process. After Hayes was convicted, the Kentucky habitual offender statute was replaced by a new statute that would have permitted him to receive, at most, an indeterminate term of 10–20 years. The new statute also set new conditions on which prior convictions could be counted. At least one of Hayes's priors did not meet these conditions. See also Tor, Avishalom. 2010. "Fairness and the Willingness to Accept Plea Bargain Offers." *Journal of Empirical Legal Studies*, 7(1):97–116. https://scholarship.law.nd.edu/cgi/viewcontent.cgi?article=1834&context=law_faculty_scholarship

[6] For life-maximum cases, 18.3% of the dispositions were by jury and 2.1% were by bench trials. For non-life-maximum cases, only 1.8% of the dispositions were by jury and 0.8% were by bench trials. Michigan Courts. 2019. 2019. "2019 Caseload Reports." re: Criminal Verdicts. https://www.courts.michigan.gov/publications/statistics-and-reports/caseload-reports/2019-caseload-reports/

[7] The number of non-habitual sentences in each offense group under each set of guidelines was:

	M2 Jud.	M2 Leg.	AWIM Jud.	AWIM Leg.	CSC Jud.	CSC Leg.	RA Jud.	RA Leg.	Total Jud	Total Leg.
Plea	951	1,218	674	828	1,493	1,422	5,987	7,523	9,105	10,991
Court	315	92	234	80	203	67	793	373	1,545	612
Jury	625	413	434	490	618	487	846	987	2,523	2,377
Total	1,891	1,723	1,342	1,398	2,314	1,976	7,626	8,883	13,173	13,980

[8] The number of habitual sentences in each offense group under each set of guidelines was:

	M2 Jud.	M2 Leg.	AWIM Jud.	AWIM Leg.	CSC Jud.	CSC Leg.	RA Jud.	RA Leg.	Total Jud.	Total Leg.
Plea	73	100	105	106	242	176	1,079	1,137	1,499	1,519
Court	19	22	13	37	32	46	95	134	159	239
Jury	53	134	70	205	140	232	196	642	459	1,213
Total	145	256	188	348	414	454	1,370	1,913	2,117	2,971

[9] Data available from the authors upon request. This does not mean that there is no relationship between conviction method and the seriousness of the defendant's prior record and or offense severity. In fact a number of relationships that were statistically significant at the p <.001 were found. For example:

- Among non-habitual murder sentences, the percentage that were pleas increased as the offense severity increased going from 58.5 to 69.5 to 78.5 for OV Levels I, II and III, respectively. This suggests that many of the second-degree murder sentences were pleas down from charges of first-degree murder.
- Among non-habitual assault sentences, 70–81% of those that scored at OV Levels I-V arose from pleas. Only 50.5% of those that scored at OV Level VI arose from pleas.
- Among non-habitual robbery sentences, the proportion that was jury-based increased steadily from OV Levels I-VI from 5.9 to 7.0 to 8.3 to 10.9 to 14.2%, then doubled to 28.6.The lower proportion of pleas among sentences scored at OV Level VI suggests that desirable plea offers were less likely to be made in these cases.
- The assault group also shows a particularly clear pattern where the proportion of pleas declined as the prior record level went up. Thus, at PRV Levels A and B the proportion of plea-based sentences was about 75%. At PRV Level C it was 62.6%. However, at PRV Levels D-F, the proportion of sentences that arose from pleas was between 55 and 57.5%.
- Robbery sentences show a similar albeit less dramatic relationship. Thus it appears that the defendants with the most serious prior records may also be less likely to receive desirable plea offers.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Chapter Ten. Systemic Factors: County of Conviction

[1] Prior studies agreed that where a defendant is sentenced is a factor in disparity; however, just how this disparity manifests is a matter of debate.

As explained in Chapter Three, the Michigan Felony Sentencing Project divided the counties into three strata based on the number of criminal dispositions in a county's circuit court, regardless of where in the state the county is located. The researchers concluded that county size was directly related to sentence length, though notably it found the longest sentences in the largest jurisdictions.

In contrast, the National Center for State Courts divided Michigan's circuits based on geography. It included smaller Washtenaw County with larger Wayne, Oakland, and Macomb Counties in a region it called SE Michigan and compared this region to all the remaining "outstate" counties, regardless of size. The NCSC researchers concluded that sentences are lower in the southeast corner of the state. That is, they found that a county's location, not its size, is related to sentence length.

While we have already observed a tendency for counties to cluster by size, with the larger jurisdictions generally having lower minimum sentences, we tested the size versus location distinction more precisely. Our results could not exactly replicate either of the prior studies because our data are only for life-maximum offenses and come from only 13 counties. But we took those counties and grouped them to simulate the opposing premises.

When we grouped the counties in the manner used by the NCSC researchers, the median for the southeast counties was eight years while the median for the outstate counties was nine years. When we placed the counties into four groups by size, with Wayne County constituting a group on its own, the medians were seven years for the larger counties (Oakland, Macomb, Kent), eight years for Wayne, 9.1 years for the medium-sized counties (Genesee, Washtenaw, Ingham, Kalamazoo), and 10 years for the smaller counties (Saginaw, Muskegon, Berrien, Calhoun, Jackson). Although the results were statistically significant at the .001 level for both approaches, grouping counties by size appears to provide more refined results.

Grouping data makes it more manageable and easier to present. However, while any grouping may be more or less accurate, all groupings smooth over differences among individual counties. Because we are examining results from only 13 of Michigan's counties, not all 83, it is possible for us to avoid groupings and assess the differences among the counties individually.

Chapter Eleven. Systemic Factors: Sentencing Judge

[1] MCR 8.111.

[2] Smith supra note 2 in ch 9 at 16–29.

Chapter Twelve. Defendant Characteristics: Sex and Age

[1] There were too few women sentenced as habitual offenders in each offense group to allow for reliable analysis.

[2] For women, 65.0% of the murder sentences and 58.3% of the assault sentences were plea-based. The comparable figures for men were 55.5% and 51.6%.

[3] Regression analyses run by sex confirm this explanation. The extent of the increase in sentence length as PRV level increased was smaller for women than for men at each level except PRV F, which included only 10 sentences for women.

[4] Frase supra note 2 in ch 6 at 102.

Chapter Thirteen. Defendant Characteristics: Race

[1] Gottschalk supra note 7 in ch 2 at 130–131.

[2] *Ibid* at 132.

[3] The reader should bear in mind that the applicable test for significance, ANOVA, is based on mean sentences, not median sentences, which is why there can be statistically significant disparities among the white and non-white assault sentences even though the medians for both are 10 years. As the note below the graph indicates, the mean assault sentence for whites was 10.5 years while the mean sentence for nonwhites was 13 years. In fact the means were both longer and more disparate than the medians for every offense group except murder.

[4] The possibility of a significant relationship between race and the number of multiple sentences was examined. There was none.

RECEIVED by MSC 4/12/2022 1:07:49 PM

[5] Frase supra note 2 in ch 6 at 116.

Chapter Fourteen. Habitual Offender Enhancements: Who Gets Chosen?

[1] CSG supra note 35 in ch 1 at 18.

[2] *Ibid* at 20–21. The CSG researchers observed that these broad disparities among counties made the cost of the habitual sentencing option both unpredictable and potentially huge.

[3] CJPC supra note 48 in ch 3. The only county that applied the enhancement more than these three was Eaton, which was not included in CSG's data or that used for the current research.

[4] Prior convictions, per se, are not included in our data. We only have prior record variable scores that indicate where the defendants had one, two, or three or more prior adult felony convictions that qualified for guidelines scoring. Therefore we do not know how many total sentences were eligible for enhancement.

[5] *Ibid.*

[6] The OMNI data had 44 sentences that received the habitual enhancement, and an additional 111 sentences that were eligible but not enhanced, that were committed by women.

Section III. Policy Implications

Chapter Fifteen. Comparing Michigan to Other Jurisdictions

[1] Lawrence, Alison. 2015. "Making Sense of Sentencing: State Systems and Policies." *National Conference of State Legislatures*. https://www.ncsl.org/documents/cj/sentencing.pdf at 5.

[2] Arizona, California, Delaware, District of Columbia, Florida, Indiana, Kansas, Minnesota, North Carolina, Ohio, Oregon, Virginia, Washington, U.S. Federal Courts.

Researchers found that states with a combination of determinate sentencing and presumptive sentencing guidelines had lower incarceration rates and smaller growth in incarceration rates than other states. Neither policy alone is enough. Moreover, "voluntary guidelines do not have the same effect. We consistently found that states with the combination of determinate sentencing and voluntary sentencing guidelines have higher incarceration rates and experienced larger growth in incarceration rates than other states. Thus, it is only by tightly controlling sentencing decisions and release decisions that states achieve lower incarceration rates and smaller growth."

Stemen, Don, Andres Rengifo, and James Wilson. 2005. *Of Fragmentation and Ferment: The Impact of State Sentencing Policies on Incarceration Rates, 1975-2002*. New York: Vera Institute of Justice. at 143.

[3] Illinois, Maine, New Mexico, New York, Wisconsin.

[4] Alabama, Alaska, Arkansas, Colorado, Louisiana, Maryland, Michigan, Mississippi, Missouri, New Jersey, Pennsylvania, Rhode Island, Tennessee, Utah.

[5] Connecticut, Georgia, Hawaii, Idaho, Iowa, Kentucky, Massachusetts, Montana, Nevada, Nebraska, New Hampshire, North Dakota, Oklahoma, South Carolina, South Dakota, Texas, Vermont, West Virginia, Wyoming.

[6] Arkansas, District of Columbia, Kansas, Maryland, Massachusetts, Michigan, Minnesota, North Carolina, Oregon, Pennsylvania, Tennessee, Utah, Washington, U.S. Federal Courts. See Robina Institute. 2021. "Sentencing Guidelines Resource Center." https://sentencing.umn.edu/

[7] The authors would like to acknowledge how helpful the website of the Robina Institute's Sentencing Guidelines Resource Center has been in providing a basic understanding of each jurisdiction's system along with citations to original sources.

[8] A mandatory minimum is a term of years below which the minimum sentence for a specific offense cannot be set regardless of the guidelines score. A statutory enhancement is an amount of time that is added to the guidelines score to increase the punishment based on a specific factor, such as use of a gun, habitual offender status, a sex offense against a child or racial bias as a motivation for the crime.

[9] This information in this section was drawn from Kansas Sentencing Commission [KSC]. 2020. "Kansas Sentencing Guidelines Desk Reference Manual 2020." https://sentencing.ks.gov/docs/default-source/2020-drm/2020-drm-text-final.pdf?sfvrsn=aac0fe3f_0 [DRM], and applicable statutes [KSA].

[10] KSA 21-6802 (2020).

[11] KSC. 2020. *Analysis of Sentencing Guidelines in Kansas: Annual Report FY 2019.* at i-ii. https://www.sentencing.ks.gov/docs/default-source/annual-reports/fy2019annualreport.pdf?sfvrsn=110bfd3f_0

[12] KSA 21-6822, 74-9101(15) (2020).

[13] KSA 21-6807(a), 21-6808(a) (2020).

[14] KSA 21-6804, 21-6805 (2020).

RECEIVED by MSC 4/12/2022 1:07:49 PM

[15] DRM supra note 9 at 30. Each grid block also indicates whether the presumptive disposition is a prison sentence or probation. The non-drug grid has several "border boxes" that allow for judicial discretion in choosing the type of disposition without having to depart from the guidelines recommendation.

[16] KSA 21-6810(d)(4) (2020); KSA 21-6811(f) (2020); DRM supra note 9 at 20

[17] KSA 21-6606(c) (2020).

[18] KSC supra note 11 at 82, Table 41.

[19] KSA 21-6804(k)(1).

[20] KSA 21-5503 (2020), KSA 21-6804(j) (2020); DRM supra note 9 at 43.

[21] KSA 21-6626 (2020); DRM supra note 9 at 40.

[22] KSA 21-6815(a) (2020).

[23] There is generally no limit on downward departures, except in cases of extreme sexual violence which may not be less than 50% of the center of the guideline range. KSA 21-6818(a) and (b).

[24] KSA 21-6819 (2020); DRM supra note 9 at 50.

[25] See KSA 21-6804 (2020), KSA 21-6805 (2020) and KSA 21-6815 (2020).

[26] KSA 21-6815(b) (2020); departures based on a plea agreement may not be appealed, KSA 21-6820(c)(2) (2020).

[27] KSA 21-6820(c) (2020); *State v. Johnson*, 286 Kan 824, 190 P3d 207 (2008).

[28] KSA 21-6820(a) and (c) (2020).

[29] KSA 21-6815(c)(2)(A) through (H) (2020).

[30] KSA 21-6815(c)(3) (2020).

[31] KSC supra note 11 at 66.

[32] The information in this section was drawn from Minnesota Sentencing Guidelines Commission. 2020. "Minnesota Sentencing Guidelines and Commentary" https://mn.gov/msgc-stat/documents/Guidelines/2020/August2020MinnSentencingGuidelinesCommentary.pdf and applicable statutes.

[33] Minn. Stat 609.2233.

[34] Minn Stat 609.229, subd 3(a).

[35] A career offender has five or more prior felonies and the present offense is part of a pattern of criminal conduct. This designation allows for an aggravated departure up to the statutory maximum. Minn. Stat 609.1095, subd 4.

[36] An engrained offender is convicted of an enumerated sex offense and meets several other criteria, including a finding that they are at great risk of reoffending without intensive psychotherapy. This designation allows a sentence that is double the presumptive term up to the statutory maximum.

[37] A dangerous offender who commits a third violent crime has two prior convictions for violent crimes and is found to be a danger to the public based on both past criminal behavior and a current aggravating factor. This designation requires a prison term that is equal to at least the length of the presumptive sentence.

[38] Minnesota Sentencing Guidelines Commission. 2021. *2021 Report to the Legislature*, at 59–65 https://mn.gov/sentencing-guidelines/assets/2021MinnSentencingGuidelinesCommReportLegislature_tcm30-463260.pdf

[39] The information in this section was drawn from: North Carolina Sentencing and Policy Advisory Commission. 2014. "Structured Sentencing Training and Reference Manual." https://sentencing.umn.edu/sites/sentencing.umn.edu/files/n.c._structured_sentencing_training_reference_manual_2014.pdf [Manual], the North Carolina Sentencing and Policy Advisory Commission [NCSPAC], and applicable General Statutes [GS].

[40] GS 15A-1340.10 (2018); Manual supra note 39 at 3.

[41] North Carolina Department of Public Safety. [NCDPS]. 2021. *Parole Process*. https://www.ncdps.gov/about-dps/boards-commissions/post-release-supervision-parole-commission/parole-process .

[42] Manual supra note 39 at 1.

[43] GS 164-37 (2020).

[44] NCSPAC. 2014. "A Citizen's Guide to Structured Sentencing." at 4. https://www.nccourts.gov/assets/inline-files/06_Citizen_Guide_to_Structured_Sentencing_2014.pdf?lT6Y2qIigjioNV0OYoxaBtN2q9JFk2mi; Manual supra note 39 at 34; credits are determined by policy of the state department of Public Safety, "Sentencing Credits", Chapter B, Sec .0100, North Carolina Public Safety, Prisons, Policy and Procedure (07/13/20), available at https://files.nc.gov/ncdps/B-.0100-07-13-20-24214150_0.pdf

[45] GS 15A-1368.2 – 1368.6 (2020); Manual supra note 39 at 36.

RECEIVED by MSC 4/12/2022 1:07:49 PM

[46] GS 15A-1340.17(c) (2020) (felony chart); GS 15A-1340.23(c) (2017) (misdemeanor chart); Manual supra note 39 at 4, 51.

[47] 15A-1340.14(b) (2020); Manual supra note 39 at 11–18.

[48] *State v. Rich*, 130 NC App 113, 502 SE2d 49 (1998); Manual supra note 39 at 12 n 5.

[49] *State v. Tucker*, 154 NC App 653, 573 SE2d 197 (2002); Manual supra note 39 at 12 n 7.

[50] Manual supra note 39 at 11.

[51] *Ibid* at 19.

[52] GS 15A-1340.16(a) to (a6) (2020).

[53] Manual supra note 39 at 20.

[54] GS 15A-1340.16(e)(7) & (15) (2020); Manual supra note 39 at 24.

[55] NCSPAC. 2019. "Structured Sentencing Statistical Report for Felonies and Misdemeanors: Fiscal Year 2018." at 18 https://www.nccourts.gov/assets/documents/publications/statisticalrpt-fy18.pdf?YTPJvi7g9H0zwNo1JCfLM6wUaI.ldf6q

[56] The information in this section was drawn from the Oregon Administrative Rules (OAR) of the Oregon Criminal Justice Commission, Chapter 213 (2021), and applicable statutes.

[57] 2020 Oregon Revised Statutes (ORS) 163.105 (2018).

[58] ORS 421.121 (reduction in term of incarceration).

[59] ORS 137.635 (determinate sentences for certain felony convictions); ORS 421.120 (reduction in term of sentence).

[60] OAR 213-017-0000 through-0011.

[61] OAR 213-018-0000 through 0090.

[62] ORS 161.535; ORS 161.605, murder in any degree under ORS 163.107 (Murder I) or 163.115 (Murder II) and treason under ORS 166.005 (Treason), which are expressly designated in the section defining the crime.

[63] ORS 164.415.

[64] OAR 213-004-0006(2).

[65] ORS 137.700(2).

[66] ORS 421.121 (reduction in term of incarceration); ORS 137.700 (listing mandatory sentence crimes).

[67] There is an exception. The statute imposing mandatory minimum prison terms for using a firearm to commit a felony permits the court to impose a less severe, guidelines-based sentence on offenders who have not previously been subject to that statute.

[68] ORS 137.667; OAR 213-008-0001.

[69] OAR 213-007-0003 through 0005.

[70] OAR 213-008-0002(1) (departure factors).

[71] OAR 213-008-0002(1)(a)(F).

[72] OAR 213-008-0002(1)(a)(F) and (2)-(3).

[73] ORS 137.123.

[74] OAR 213-012-0020(2)(b).

[75] The information in this section was drawn from Utah Sentencing Commission. 2020. "2020 Utah Adult Sentencing & Release Guidelines." https://justice.utah.gov/wp-content/uploads/2020-Adult-Sentencing-and-Release-Guidelines.pdf [Guidelines], Utah Sentencing Commission, and applicable statutes.

[76] Utah Code 77-27-5.4.

[77] Guidelines supra note 75 at 22.

[78] *Ibid.*

[79] Utah Code 76-3-401(1); Guidelines supra note 75 at 17–18.

[80] Guidelines supra note 75 at 18.

[81] United States Sentencing Commission. 2018. "Guidelines Manual 2018." https://www.ussc.gov/sites/default/files/pdf/guidelines-manual/2018/GLMFull.pdf [Manual].

[82] USSC. 2018. "Federal Sentencing: The Basics." at 1. https://sentencing.umn.edu/sites/sentencing.umn.edu/files/2019_us_federal_sentencing_the_basics_primer.pdf

[83] *Peugh v. United States*, 569 US 530, 535 (2013).

[84] S Rep No 98-225 at 38 (1983), reprinted in 1984 USCCAN, 3182. The Senate Judiciary Committee report is the primary legislative history of the SRA. See *Mistretta v. United States*, 488 US 361, 366 (1988).

[85] See *Dorsey v. United States*, 567 US 260, 265 (2012).

[86] See 28 USC 991(a).

[87] USSC. N.d. "Organization." https://www.ussc.gov/about/who-we-are/organization

[88] *Peugh*, 569 US at 544 (2013).

RECEIVED by MSC 4/12/2022 1:07:49 PM

[89] *Gall v. United States,* 552 U.S. 38, 49 (2007).

[90] *Kimbrough v. United States*, 552 US 85, 101 (2007) (citing *Booker*, 543 US at 245-26)

[91] See, e.g., *Molina-Martinez v. United States*, 136 S Ct at 1345 (2016) (citing *Gall*).

[92] See *United States v. Nelson*, 237 Fed. Appx. 819, 821 (2007) (per curiam).

[93] See, eg, *Nelson v. United States*, 555 US 350, 351 (2009); *Rita v. United States*, 551 US 338, 351 (2007) ("[T]he sentencing court does not enjoy the benefit of a legal presumption that the Guidelines sentence should apply"); *United States v. Perez-Perez*, 512 F.3d 514, 516 (9th Cir. 2008).

[94] *Gall* at 41; *Peugh v. United States*, 569 US 530 (2013).

[95] *Peugh v. United States*, 569 US 530, 537 (2013).

[96] *Kimbrough v. US*, 552 U.S. 85 at 109.

[97] 18 USC § 3582(c)(1) (2019); the First Step Act of 2018, 115 PL 391§ 603(b)(1), amended 18 USC § 3624(b) to provide that such reductions can sometimes be made on motion of the defendant (a motion by the Director of the Bureau of Prisons is no longer essential).

[98] 552 US 38, 49 (2007) (the district court should begin all sentencing proceedings by correctly calculating the applicable guideline range, and "to secure nationwide consistency, the Guidelines should be the starting point and the initial benchmark."); see also USSG §1B1.1(a)–(c) (Application Instructions).

[99] *Ibid*; 18 USC § 3553(a)(4).

[100] See USSG §1B1.1(c); see also *United States v. Hughes*, 401 F3d 540, 546 (4th Cir 2005); *United States v. Stone*, 432 F3d 651, 655 (6th Cir 2005). [This paragraph is from USSC. 2020. "Primer on Departures and Variances." at 2 https://www.ussc.gov/sites/default/files/pdf/training/primers/2020_Primer_Departure_Variance.pdf

[101] 28 U.S.C. § 994(b) (2019) ("if the minimum term of the range is 30 years or more, the maximum may be life imprisonment.").

[102] USSG § 2A1.2.

[103] USSG § 3A1.1.

[104] USSG §1B1.3, comment; USSC. 2020. "Primer on Relevant Conduct" at 2 https://www.ussc.gov/sites/default/files/pdf/training/primers/2020_Primer_Relevant_Conduct.pdf

[105] USSC supra note 104 at 1

[106] USSG § 3.D.1.

[107] USSG §§ 4A1.1-.2.

[108] USSG §§ 4A1.3; 4B1.

[109] USSG §§ 4A1.3; 4B1.

[110] USSG § 4A1.2(a)(2).

[111] 18 USC 3559(c) requires life imprisonment on conviction in federal court of a "serious violent felony" and has two or more prior convictions in federal or state courts, at least one of which is a "serious violent felony."

[112] USSG § 5K2.21.

[113] USSC at 13, https://www.ussc.gov/guidelines/primers/departures-and-variances.

[114] USSC supra note 100 at 43 (nns omitted)

[115] See *United States v. Vizcarra*, 668 F3d 516 (7th Cir 2012); USSG § 1B1.1 cmt n4; however, in *United States v. Fugate*, 964 F.3d 580 (6th Cir 2020) impermissible double-counting was found for being subjected to dual four-level enhancements under USSG § 2K2.1(b): one under 2K2.1(b)(5) for trafficking in firearms, and the other under § 2K2.1(b)(6)(B) for possessing or trafficking firearms.

[116] 18 USC 922(c).

[117] USSC. 2020. "Statistical Information Packet: Fiscal Year 2020, State of Michigan." At 12, Table 8. The report comparing US Courts nationally to Michigan federal courts is at https://www.ussc.gov/sites/default/files/pdf/research-and-publications/federal-sentencing-statistics/state-district-circuit/2020/mi20.pdf The main set of data is at https://www.ussc.gov/research/data-reports/geography/2020-federal-sentencing-statistics.

[118] 543 U.S.220 (2005).

[119] 498 Mich 358.

[120] MCL 750.529.

[121] MCL 750.520b.

[122] MCL 28.293(3).

[123] MCL 750.321.

[124] MCL 750.357.

RECEIVED by MSC 4/12/2022 1:07:49 PM

[125] MCL 333.7403(2)(b)(i). Although as discussed in Chapter One, the Legislature has moved a number of crimes from their original grids into higher grids, thereby increasing the cell ranges, without any changes in the maximum penalties.

[126] MCL 777.57(2)(a).

[127] USSG §1B1.3(a).

[128] USSG § 5K2.21.

[129] The American Law Institute supra note 22 in ch 6 at s6B.06(2)(b) and., s7.03(2)(b). See also American Bar Association. N.d. "Criminal Justice Standards Sentencing." at Standard 18-3.6 https://www.americanbar.org/groups/criminal_justice/publications/criminal_justice_section_archive/crimjust_standards_sentencing_toc/ ; and Reitz, Kevin R. 1993. "Sentencing Facts: Travesties of Real-Offense Sentencing." *Stanford Law Review*, 45(3):523–573. https://doi.org/10.2307/1229007

[130] The American Law Institute supra note 22 in ch 6 at s6B.06(3) and s7.03(3).

[131] Guidelines supra note 75 at p 22.

[132] KSA 2020 Supp 21-6815(c)(3).

[133] OAR 213-008-0002 (2020).

[134] Pennsylvania Commission on Sentencing. 2012. "Sentencing Guidelines Implementation Manual: 7th Edition." at 226. https://sentencing.umn.edu/sites/sentencing.umn.edu/files/Pennsylvania%20Sentencing%20Guidelines%20Manual%207th%20Edition.pdf

[135] SENTAC. 2019. "Delaware Sentencing Accountability Commission Benchbook." at 132-136. https://cjc.delaware.gov/wp-content/uploads/sites/61/2019/01/Benchbook-2019.pdf

[136] Mich. Comp. Laws § 769.34(3)(b) (2016). The viability of this provision post-Lockridge is unclear.

Chapter Sixteen. Should the Guidelines Be Enforced?

[1] 570 U.S. 99 (2013).

[2] Chief Justice Roberts, in a dissent joined by Justices Scalia and Kennedy, expressed the view that the jury trial right was not implicated because the jury's verdict authorized the judge to impose any sentence between the mandatory minimum and the statutory maximum. Judges have long held the discretion to consider sentencing factors they deemed relevant that were not considered by the jury. No one saw a problem if the trial judge, on his own, decided to raise a minimum sentence by two years because the defendant used a gun. Rather, Justice Roberts said, that by prohibiting a legislative mandate from requiring the judge to impose a particular sentence, the majority was really protecting the discretionary power of sentencing judges.

[3] 498 Mich 358 (2015).

[4] What have been traditionally considered to be "mandatory minimums" in Michigan are penalties like the one in *Alleyne,* which are prescribed by statute for certain crimes. As noted in Chapter One, significant examples are the 25-year minimums required for first-degree CSC with a child under 13 and for certain felonies when the defendant is being sentenced as a fourth habitual offender. The guidelines, by contrast, do not require the imposition of any particular sentence. And the breadth of the cell ranges, especially for the life-maximum offenses, preserves an enormous amount of judicial discretion. Nevertheless, departures occur routinely. Moreover the guidelines enabling legislation expressly recognizes a mandatory minimum as being something different from a guidelines minimum. It provides that a mandatory minimum trumps a guidelines recommendation and states: "Imposing a mandatory minimum sentence is not a departure under this section." MCL 769.34(2)(a). However, the *Lockridge* majority considered and rejected the argument that Michigan's guidelines did not set a mandatory minimum for constitutional purposes.

[5] 498 Mich 358, 461 (2015).

[6] 500 Mich 453, 482-484 (2017).

[7] See, e.g., *People v. Pohl*, 445 Mich 915 (1994) (minimum sentence of 86 months for breaking and entering was seven times the guidelines maximum-minimum of 12 months); *People v. Hughes,* unpublished per curiam opinion of the Court of Appeals issued 6/26/98 (Docket # 191776) (40 year minimum for assault with intent to murder was excessive departure from guidelines range of 3-8 years); *People v. Cramer,* 201 Mich App 590 (1993) (60 year minimum for criminal sexual conduct was unwarranted departure from guidelines range of 10–25 years).

[8] See *People v. Merriweather*, 447 Mich 799 (1994) (majority approved a sentence of 60-120 years that was purposefully designed to prevent a defendant who brutally raped and tortured an elderly woman from ever being eligible for parole); *People v. Houston*, 448 Mich 312 (1995) (approved a sentence of 25–50 years for first-degree

RECEIVED by MSC 4/12/2022 1:07:49 PM

criminal sexual conduct despite a guidelines range of six to 10 years); *People v. Lemons*, 454 Mich 234 (1997) (sentences of both life and 60–90 years were found proportional for a 45-year-old defendant convicted of criminal sexual conduct); *People v. Hansford (After Remand),* 454 Mich 320 (1997) (a 40–60 year sentence for a habitual offender convicted of receiving and concealing stolen property was justified by the serious nature of the crime, the defendant's seven prior felonies and his clear inability to reform); *People v. Montney,* Court of Appeals Docket #151454, 8/5/1994, unpublished memorandum, *lv den (*defendant with no prior record was sentenced to 50–75 years for his participation in a second-degree murder at the age of 16, although the guidelines range was 10–25 years).

[9] *People v. Phillips (After Second Remand)* 227 Mich App 28 (1997)(65–150 year sentence for murder of off-duty state trooper upheld).

[10] *People v. Leon Orlando Echols,* order of Supreme Court denying leave to appeal issued March 7,2014 (Docket No. 147533) (Cavanagh, J. dissenting). Justice Cavanagh also found that the sentence was designed to avoid the possibility of parole that would exist with a life sentence.

[11] 469 Mich 247 (2003). The *Babcock* Court stressed that the principle of proportionality remained relevant as the standard against which the decision to depart was to be assessed. In establishing the sentencing guidelines, the Legislature was subscribing to the "deeply rooted" belief that the punishment should fit both the crime and the criminal. The scoring of offense and prior record variables reflects this belief. Therefore, a departure must contribute to making a sentence more proportionate than a sentence within the guidelines range.

[12] 482 Mich 292 (2008).

[13] MCL 769.34(3) now states: 3) A court may depart from the appropriate sentence range established under the sentencing guidelines set forth in chapter XVII if the departure is reasonable and the court states on the record the reasons for departure. All of the following apply to a departure:

(a) The court shall not use an individual's gender, race, ethnicity, alienage, national origin, legal occupation, lack of employment, representation by appointed legal counsel, representation by retained legal counsel, appearance in propria persona, or religion to depart from the appropriate sentence range.

(b) The court shall not base a departure on an offense characteristic or offender characteristic already taken into account in determining the appropriate sentence range unless the court finds from the facts contained in the court record, including the presentence investigation report, that the characteristic has been given inadequate or disproportionate weight.

[14] 500 Mich 453 (2017).

[15] *People v. Dixon-Bey*, 321 Mich App 490, 525 (2017).

[16] There is also some irony in the fact that departures can be based on considerations the Sentencing Guidelines Advisory Committee expressly rejected as inappropriate for inclusion in the judicial sentencing guidelines. See text at note 14, *supra.*

[17] Note 14, *supra.*

[18] Note that the Michigan Supreme Court denied leave to appeal in *Mead, Tucker, Scott, Odom, Schurz and Latimer.* An application for leave to appeal filed in *Boak* on July 27, 2021 was pending when this report went to press.

[19] Unpublished opinion of Court of Appeals issued Mar. 19, 2019 (Docket No. 341504).

[20] Unpublished opinion of Court of Appeals issued Mar. 5, 2019 (Docket No. 336563).

[21] Unpublished per curiam opinion of the Court of Appeals issued Nov.21, 2019 (Docket No. 343305).

[22] Unpublished per curiam opinion of the Court of Appeals issued June 3, 2021 (Docket No. 340201).

[23] Unpublished per curiam opinion of Court of Appeals issued Aug. 29, 2019 (Docket No. 341688).

[24] Unpublished per curiam opinion of Court of Appeals issued Aug. 29, 2019 (Docket No. 343351).

[25] Unpublished per curiam opinion of the Court of Appeals issued July 11, 2019 (Docket No. 339325)

[26] 327 Mich App 297 (2019)

[27] Unpublished per curiam opinion of the Court of Appeals issued Aug. 27, 2020 (Docket No. 340420).

[28] Unpublished per curiam opinion of the Court of Appeals issued March 21, 2019 (Docket No. 336692).

[29] In *People v. Beck,* 504 Mich 605 (2019), the jury acquitted the defendant of open murder but convicted him of being a fourth offender felon in possession of a firearm. The judge based a departure of 20 years on its finding by a preponderance of the evidence that the defendant had committed the murder. The Supreme Court held that due process bars a court from basing a sentence on its own finding that a defendant had committed conduct of which he was acquitted. There is no constitutional prohibition on a judicial finding of guilt of conduct the jury never considered, such as charges the prosecution chose not to bring.

RECEIVED by MSC 4/12/2022 1:07:49 PM

APPENDICES

RECEIVED by MSC 4/12/2022 1:07:49 PM

TECHNICAL APPENDIX A. THE DATA AND METHODS OF ANALYSIS

As explained above, by focusing on the four most common life-maximum offenses, the instant research examines the effectiveness of the sentencing guidelines in promoting proportionality and decreasing disparities in the cases where judges retain by far the greatest amount of sentencing discretion. These are the offenses for which guidelines ranges are the broadest, recommended minimums are the longest, and the sentences imposed have the greatest impact on both individuals and the prisoner population.

We have built on, expanded, and refined the focus of prior research. We attempt to measure the extent to which factors scored in the legislative guidelines explain the length of prison sentences as well as the extent to which guidelines recommendations are actually followed. We then examine the impact on sentence length of systemic factors including county size, conviction method, and the identity of the sentencing judge, and of defendant characteristics like age, gender, and especially race. Throughout, we analyze each offense group separately to account for the fact that the nature of the crime is, in itself, an exceedingly important factor in determining the sentence and that lumping all the life-maximum offenses together masks the substantial differences among them. We also pay close attention to a factor that became particularly relevant with the enactment of the legislative guidelines—the prosecution's decision to charge the defendant as a habitual offender.

A. Data

The data for our research come from two MDOC databases. The first is the Correctional Management Information System (CMIS), which contains all the information necessary to operate a prison system, including not only basic sentencing information but also the details of institutional life, like transfers between prisons, program participation, and disciplinary citations. CMIS includes sentences from all three guidelines time periods, allowing for the comparison of sentences over time. The second database, the Offender Management Network Information system (OMNI), is smaller but has more details about each sentence, including guidelines scoring. OMNI only includes sentences from the legislative guidelines period. All the sentences in OMNI are also in CMIS. We obtained these databases in early 2013. Thus, our dataset includes cases in which the offense was committed by December 31, 2012, and sentencing information was available.

Which set of sentencing guidelines applies in a given case depends on the date the offense took place. To accurately analyze differences among the guidelines periods, we made several decisions about defining time periods. For the CMIS data, to examine the first judicial guidelines period, we use data with offenses from 1984 to1988. We excluded data from 1983, as it was the first year of the guidelines and reporting was inconsistent. The second judicial guidelines period has data from the entire period, 1989–1998, and the legislative data has offenses committed from 1999 to 2012. The available OMNI data includes sentences imposed from 2003–2012 for offenses committed from 1999–2012.

The data from the first judicial guidelines are utilized only in Chapter Five to examine changes in the length, nature, and number of sentences over time. Because the guidelines were changed substantially after 1988, and the second judicial period is more directly comparable to the legislative period, only the second judicial guidelines are used for comparison in later chapters. To optimize the amount of data available regarding the legislative guidelines, we utilize CMIS data whenever possible. The exception is when we are examining elements of the sentencing guidelines, such as scores for offense and prior record variables or location of the sentence within the applicable range, which are available only in OMNI. Throughout the report, we indicate below tables and graphs which data source was used.

The unit of analysis is the sentence, not the individual defendant. The research goal is to explain, to the extent possible, what factors influence how judges exercise their discretion to impose a sentence of "life or any term." A key focus is on how sentencing guidelines recommendations affect that decision. Therefore, we do not attempt to measure other factors that help determine how long the person must remain in prison, such as whether they have additional sentences for other offenses or whether multiple sentences are imposed consecutively.[1]

RECEIVED by MSC 4/12/2022 1:07:49 PM

We count every non-redundant sentence imposed for the most common life-maximum offenses committed in each guidelines period. Redundant sentences share all the following characteristics:

- identical minimum sentences
- same statutory citation for offense
- same offense date
- same county
- same sentencing date

Redundant sentences are excluded because they risk artificially skewing the results. Thus, for instance, if someone is convicted of four counts of armed robbery because there were four victims but the sentence for each count is seven to 20 years, that sentence is only counted once. Similarly, if a defendant convicted of criminal sexual conduct receives identical sentences for each of multiple penetrations of the same victim on the same day, that sentence is only counted once. It is assumed that the same guidelines scoring and the same judicial reasoning produced the identical outcomes and that there is nothing to be gained from including redundant sentences in the analysis. On the other hand, counting the same sentence multiple times would distort the calculation of mean and median sentences for all the robbery or all the CSC sentences imposed in a given time period or a given county. The number of redundant sentences that were eliminated was 319 for the first judicial period and 1,190 for the second judicial period. For the legislative period, 4,013 were eliminated from the CMIS data and 2,797 were eliminated from the OMNI data.

Sentences for different life-maximum offenses occurring in the same incident are counted, such as if the defendant was convicted of robbing and assaulting the same victim. Multiple sentences are also counted if the defendant committed the same crime on different dates or on the same date in different counties (e.g., by robbing a series of gas stations in the same county in the course of a week or in different counties on the same day). And, of course, if a defendant received prison sentences for life-maximum offenses in more than one guidelines period, all of that person's non-redundant sentences would be counted. This means that the same individual can and does appear in our data multiple times; specifically, 827 individuals showed up in two time periods and 22 had sentences in all three time periods. The frequency of multiple sentences is discussed in Chapter Five.

In our primary analyses we only include sentences that have a minimum and maximum term of years. Parolable life sentences are excluded because they have no minimum terms that can be quantified and compared statistically to the rest of the data. This filter resulted in the exclusion of 360 sentences from the first judicial period, 713 sentences from the second judicial period, and 312 sentences from the legislative period. In Chapter Five, we briefly examine the declining use of discretionary life sentences over the three time periods.

The composition of our entire dataset, after excluding redundant sentences and parolable life terms, is reflected in Table A1. The total number for each guidelines period is far greater than the number of sentences for life-maximum offenses examined by any previous research.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table A1. Number of Sentences by Guidelines Period, Offense Type, and Habitual Offender Status					
	1st Judicial 1984–1988 CMIS	2nd Judicial 1989–1998 CMIS	Legislative 1999–2012 CMIS	All Periods CMIS TOTAL	Legislative 2003–2012 OMNI
Second-degree Murder (M2)	**854**	**2,036**	**1,979**	**4,869**	**1,399**
Non-habitual	834	1,891	1,723	4,447	1,218
Habitual	20	145	256	422	181
Assault with Intent to Murder (AWIM)	**587**	**1,530**	**1,746**	**3,863**	**1,089**
Non-habitual	548	1,342	1,398	3,287	884
Habitual	39	188	348	576	205
First-degree Criminal Sexual Conduct (CSC1)	**1,332**	**2,728**	**2,430**	**6,490**	**1,510**
Non-habitual	1,226	2,314	1,976	5,516	1,225
Habitual	106	414	454	974	285
Armed Robbery (RA)	**3,637**	**8,996**	**10,796**	**23,429**	**6,886**
Non-habitual	3,318	7,626	8,883	19,827	5,659
Habitual	319	1,370	1,913	3,602	1,227
TOTAL	**6,410**	**15,290**	**16,951**	**38,651**	**10,884**
Non-habitual	5,926	13,173	13,980	33,079	8,986
Habitual	484	2,117	2,971	5,572	1,898

The four most common life-maximum offenses are second-degree murder, assault with intent to murder, first-degree criminal sexual conduct, and armed robbery. As we grouped the sentences by offense type, we included, where applicable, highly similar crimes that also carry a maximum sentence of life or any term.[2] The data come from 13 Michigan counties:

-Berrien -Calhoun -Genesee -Ingham -Jackson -Kalamazoo -Kent
-Macomb -Muskegon -Oakland -Saginaw -Washtenaw -Wayne

These were the counties with the largest number of commitments to prison for assaultive offenses in 2013. Collectively, they comprised 75% of the total.[3] Since one of our key foci is county differences in sentence length, it was important to include only counties with a sufficient number of cases to allow for meaningful comparisons.[4]

B. Analytic Strategies

1. Univariate Analysis: Examining Sentence Lengths

The first step in our analyses is to examine how the minimum sentence length for each offense group changed in each guidelines period. To do this, we report the median length of sentences. It is common

RECEIVED by MSC 4/12/2022 1:07:49 PM

practice to report the mean or average, which is calculated by adding together all the values and dividing by the total number of values. However, this measure of central tendency is sensitive to extreme values or outliers—cases at either the high or low end that are not representative of most cases but can pull the average in one direction or the other. A sufficient number of outliers will cause the data to be skewed.

In contrast, the median is defined as the midpoint in a set of values. To determine the median, all values are listed in order, and the value in the middle is the median. Imagine, for instance, 101 sentences listed from shortest to longest. The sentence in the 51st position is the median.

If the distribution of cases creates a bell-shaped curve, with the majority clustering toward the center and fewer cases at each end, the mean and the median will tend to be quite similar. If the cases tend to cluster more at either the lower or higher end of the scale, the mean and median will be quite different.

Michigan sentencing data are highly skewed. They contain extreme values (i.e., decades-long minimums that affect the mean), resulting in an "average" sentence that is not representative of most sentences. The bulk of the cases are at the lower end of the range of sentence lengths. However, the cases with very high sentences pull the average up substantially.[5] By utilizing the median for our reporting, we can have a more accurate idea of how long most sentences are, as medians are not impacted by extreme scores.

In considering the reported median sentence lengths, the reader should bear in mind what the medians can and cannot tell us. Neither the median nor the mean is necessarily illustrative of the actual sentences people received. While half of all sentences were above the median and half were below, it may be that few if any were actually at the median. Thus, the median does not tell us what sentence in a group was most common.

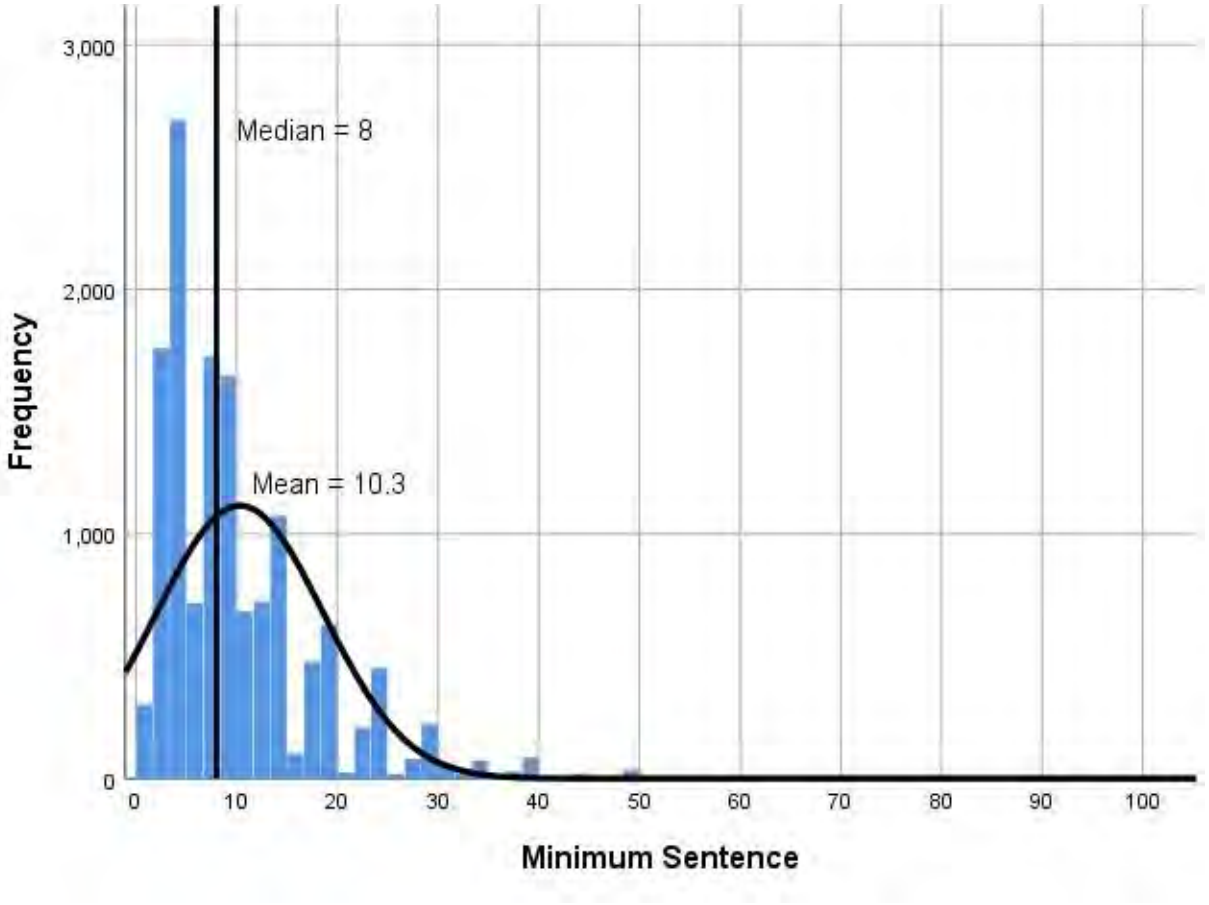
What the median does give us is a consistent way of comparing sentences across groups or over time. So, for instance, if we know the median sentence for murder under the judicial and legislative guidelines, we can see if the sentences grew longer. If we know the median sentence for murder in each of 13 counties in the legislative period, we can see if the application of the guidelines produced similar sentences around the state. Since the purpose of the research is to compare relative sentence lengths depending on various factors, having accurate medians to represent sentence length accomplishes that goal. For the reader's information, Table C1 in the Appendix provides information about the lowest and highest sentences for each offense group, by county, during the legislative period as well as the range of sentences between the 25th and 75th quartiles.

Figure A1 illustrates this pattern for everyone not sentenced as a habitual offender under the legislative guidelines. The vertical line bisects the horizontal axis at exactly the halfway point. This results in a median of eight years. The peak of the curve is the mean sentence of 10.3 years. Fifty percent of the minimum sentences are concentrated in a range between four and 14 years. Another 25% are at the low end, with sentences below four years. But the other end of the curve is a long tail, with 25% of the sentences—almost 3,500—extending past 14 years to 20, 25, 30, 40, and even 50 years. In fact, the longest minimum sentence is 100 years.

Both the median and the mean for the total group are relatively low because more than 60% of the cases are armed robberies, which have lower sentences than the other life-maximum offenses. But the pattern repeats for nearly every offense group in each guidelines period.[6] Generally, the mean is anywhere from one to five years and seven to 50% longer than the median.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Figure A1. Frequency of Non-Habitual Minimum Sentences in Legislative Period



CMIS data
Using the mean is necessary for some additional computations that we describe next, specifically ANOVA and regression analyses, and so we employ it for those purposes.

2. Bivariate Analysis: The Relationship Between Sentence Length and Individual Variables

The second step of our analyses is to see how sentence length varies in relation to other variables, such as sentencing guidelines scores or county of conviction. To do so, we utilize two forms of bivariate analysis, cross-tabulations and means comparison. Cross-tabulations show the frequency of one variable in relation to two or more categorical variables; "categorical" meaning they are grouped into categories that describe their distinguishing characteristics, such as guilty plea versus jury trial or small, medium, and large counties. For example, we can examine how many sentences of each offense happened in each county, or were resolved through the different methods of conviction.

The means comparison also examines how two, or more, variables vary in relation to each other. In this analysis, the dependent variable is continuous (i.e., when one of the variables is a number, like minimum sentence length, and the independent variable[s] is categorical). For example, we can examine differences in sentence length depending on the county and the method of conviction. While cross-tabulations report the frequency of raw data, the means comparison summarizes those frequencies into means for statistical comparison.

RECEIVED by MSC 4/12/2022 1:07:49 PM

To evaluate if these differences are meaningful we utilize tests of statistical significance that assess the likelihood that an observed bivariate relationship differs significantly from what could have occurred by chance. The chi-square statistic is used when the variables in question are categorical. The other analysis, ANOVA, "a test for the difference between two or more means," is used when the dependent variable is continuous or ratio-level data.[7]

The significance of all of the findings is reported in the relevant tables. Following convention, the level of significance is displayed at the bottom of each table or figure as a p-value. To say that something is statistically significant at the $p < .05$ level means that there is only a 5% likelihood that the relationship occurred by chance. At $p < .01$ the likelihood of a chance relationship falls to 1%. A p-value of $< .001$ means that we are 99.9% sure the relationship is not occurring by chance.

3. Multivariate Analysis: How Multiple Factors Interact to Produce a Sentence

To understand the extent to which multiple variables each helps determine sentence length we use multiple regression to conduct causal analysis. Multiple or linear regression is one of the most widely used methods of analysis in the social sciences. It simultaneously examines various factors that might cause the dependent variable (in our case, sentence length) to change in order to examine the unique contribution of each variable.[8] This is a progression from the cross-tabulations examined in our bivariate analyses because it allows for the interpretation of one variable while "controlling" for others. For example, using multiple regression we can understand how much of the difference in sentence length is due to factors like offense type or sentencing guidelines scores while accounting for the impact of other factors like county size and whether the conviction resulted from a plea or a trial. The multivariate analyses is explained in more detail in Appendix B.

---

[1] If the research focus was factors that drive mass incarceration, we would count people not sentences. No matter how many sentences a defendant receives, the person occupies one prison bed. To measure the shortest time the person will stay in prison requires identifying the "controlling sentence," whether that is a life term or the longest minimum of an indeterminate term. The minimum term(s) of any consecutive sentence(s) must be added to the controlling sentence, and any available good time or disciplinary credits as well as credit for time spent in jail prior to sentencing must be subtracted. Even the minimum length of stay does not, of course, determine an individual's actual length of stay, since that ultimately depends (except in cases of mandatory life without parole) on the decision of the parole board to release the person.

We recognize that the fact that a judge may choose to or be required to impose multiple sentences consecutively may influence the judge's selection of the minimum sentence for each offense. However, there is no way to determine from quantitative data whether or to what extent that influence occurred. Even more to the point, since appellate courts assess the proportionality of each sentence separately—without regard to their aggregate length—our empirical analysis of each sentence independently, regardless of whether they are served concurrently or consecutively, accords with Michigan law. See *People v. Miles* supra note 52 in ch 1

[2] The Second Degree Murder (M2) group includes only MCL 750.317, Second Degree Murder.

The Assault with Intent to Murder (AWIM) group includes:

MCL 750.83	Assault with intent to commit murder
750.91	Attempt to commit murder
750.157b	Solicitation to commit murder
750.204d	Sending explosives with intent to frighten, injure or kill a person
750.207	Placing explosive substance with intent to frighten, injure or kill a person
750.317a	Delivery of schedule 1 or 2 controlled substance causing death
750.90a	Intentional conduct causing a miscarriage or stillbirth.

First Degree Criminal Sexual Conduct (CSC1) includes the eight variations of First Degree Criminal Sexual conduct defined in MCL 750.520b(1)(a)-(h).

The Armed Robbery (RA) group includes:

MCL 750.529	Armed robbery
750.529a	Carjacking
750.89	Assault with intent to rob and steal.

RECEIVED by MSC 4/12/2022 1:07:49 PM

[3] Michigan Department of Corrections. 2014. Michigan Department of Corrections 2013 Statistical Report. Table A1c at pp. A-10. https://www.michigan.gov/documents/corrections/2014-04-04_-_MDOC_2013_Statistical_Report_-_Vers_1_0_452815_7.pdf

[4] Note that, until 1997, the City of Detroit, which is in Wayne County, had its own criminal court. In 1997, the Recorder's Court for the City of Detroit was merged with the Wayne County Circuit Court, and Recorder's Court judges became members of the Wayne Circuit Court bench. The sentences included in our data for the first and second judicial periods were adjudicated in either Recorder's Court or the Wayne County Circuit Court, although they were all prosecuted by the Wayne County Prosecutor's Office. We treat them all as Wayne County sentences for purposes of our data analysis.

[5] If the data skewed to the right so that the majority of cases were at the high end and the extreme sentences were much shorter than normal, the median would be longer than the mean.

[6] The sole exception is the 172 murder cases during the judicial guidelines period for which people were sentenced as habitual offenders.

[7] The decision to test for significance using ANOVA is an imperfect solution. ANOVA is means-based and, for the reasons explained, we are reporting medians. We explored available tests for examining the significance of medians, specifically Independent Medians and Mann Whitney U. We did not choose either of these tests because the distribution of sentences within between our comparison

[8] Allison, Paul. 1999. *Multiple Regression: A Primer.* Pine Forge Press. Pp. 3

RECEIVED by MSC 4/12/2022 1:07:49 PM

TECHNICAL APPENDIX B. DETERMING SENTENCE LENGTH: WHICH FACTORS COUNT THE MOST

The bivariate analyses identified multiple factors that have a statistically significant relationship to sentence length. We can say with a high degree of certainty that each of these factors has an impact on the length of the minimum sentence. But how much of an impact does each one have? And does the extent of the impact change depending on the presence or absence of other significant factors? For instance, is the defendant's prior record level more important if the offense is robbery or if it is murder? How much difference does it make if the defendant was sentenced as a second offender or a fourth offender? How much of the difference in sentence length for two defendants with similar prior records can be explained by the county of conviction or whether one of them pleaded guilty?

To answer such questions, we use a statistical tool called multiple linear regression. Multiple linear regression or multiple regression is one of the most widely used methods of statistical analysis in the social sciences. The popularity of this method stems from its ability to simultaneously examine various factors that change the dependent variable (in our case sentence, length), in order to examine the unique contribution of each variable.[1] This analysis is a progression from the cross-tabulations examined in the bivariate analyses because it allows for the interpretation of the impact of multiple independent variables at once, while "controlling" for others.

Regressions: Explanation and Estimation

Regressions can perform two functions. First, they can identify how much individual independent variables or groups of independent variables explain the variance in the dependent variable. For this research we are examining the effect of the guidelines factors, systemic factors and defendant characteristics on the dependent variable of sentence length. This aspect of regression analysis allows us to see the relative importance of factors that we know have a significant relationship to sentence length when they are combined with other significant factors. Results can show us, for instance, if the importance of prior record changes depending on the offense or if the importance of a habitual offender enhancement changes depending on the county of conviction.

Regression analysis also allows us to estimate how a baseline sentence will change if particular variables are adjusted. For instance, if the defendant convicted of an armed robbery received a five-year sentence when his prior record level was A, by what percentage would that sentence increase if his prior record level became C? Or if he received five years when he had no habitual offender enhancement, how much more would the sentence increase if he were enhanced to fourth offender status? These predictions or estimated sentences can be tailored to include many details, including the county of conviction and whether the defendant chose a guilty plea or a jury trial.

The Data and Variables

For our regression analyses, we utilize the OMNI data described in Appendix A, as it includes the guidelines factors to be used as independent variables.[2]

The dependent variable in our analyses is the natural logarithm (ln) of minimum sentence in months. Utilizing the natural logarithm achieves two important goals. First it is a technique to normalize skewed data, which is required for using multiple regression. Second, this approach follows prior research that suggests using the natural logarithm of the dependent variable of sentence length is more accurate.[3] In combination with the use of categorical independent variables, the regressions can be interpreted as a percentage increase in sentence length.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Independent Variables

For our independent variables, we utilize categorical variables, which have been recoded into dichotomous variables, to represent whether a characteristic is present. Our independent variables can be divided into three groups: guidelines factors, systemic factors, and defendant characteristics, following the prior research examining sentencing guidelines in Michigan. The guidelines factors include offense type, offense level, prior record level, habitual offender status, and departure status. Systemic factors include conviction method and county of conviction. The last group of defendant characteristics includes age, race, and sex. We expand upon the prior research in three important ways. First, we separate the data by offense type, while prior research has only examined at the grid level. Second, we conduct separate analyses based on habitual offender status. And third, we conduct separate analyses by county.

Utilizing dummy variables assists in meaningful interpretation of our categorical independent variables. When utilizing this type of variable, one of the categories has to be excluded and is used as the reference category for interpretation, as displayed in Table B1. For example, there are six prior record levels. In our regressions, we omit PRV level A to use it as the reference category, so the included PRV levels can be interpreted as a percentage increase from PRV level 1. The independent variable groups, the reference categories, and which variables are included in the regressions appear in the table below.

Table B1. Independent Variables in the Regression Analyses

Variable Group	Reference Category	Variables included in Regression
Offense	RA	M2, AWIM & CSC
OV Level	OV Level 1	OV Levels 2–6
PRV Level	PRV Level 1	PRV Levels 2–6
Habitual Status	Non-Habitual	Habitual levels 2–4
Departure	In Range	Departure below, departure above
Conviction Method	Plea	Court, Jury
County	Wayne County	Berrien, Calhoun, Genesee, Ingham, Jackson, Kalamazoo, Kent, Macomb, Muskegon, Oakland, Saginaw, and Washtenaw
Defendant Characteristics		
Age	Ages 20-29	Under 17, Age 17, 18–19, 30–39, and 40 and older
Race	Coded as Non-White= 0 and White= 1	
Sex	Coded as Male =0 and Female =1	

Analytical Strategy

To best understand the importance of each of these variables, we utilize sequential, or hierarchical, multiple regression, which is used to determine the variance explained by different sets of independent variables.[4] Since the order in which you enter independent variables influences the outcome, we followed the logic of the Michigan sentencing guidelines and the relationships examined so far in this research to determine the order of our variables. The offense to be sentenced is the primary determining factor in sentence length, as it determines what grid will be used. This is followed by the offense and prior record level, and habitual status. Next are upward and downward departures, which can be viewed as a rejection of the sentencing guidelines that have been shown to be a prevalent and consistent contributor to sentence length. Then come the method and county of conviction, which are not directly related to the guidelines but have been shown to be important in examining sentence length. The last variables entered are the

RECEIVED by MSC 4/12/2022 1:07:49 PM

demographics or defendant characteristics, which theoretically should not have an influence on sentence length and, if they do, could be an indication of discrimination.

The following sections of this chapter will proceed with explanations of how to read and interpret the explanation and estimation models, using the full model as examples. The full model includes all of the offenses and counties as independent variables. The offense-specific and county-specific analysis tables are included here but interpreted in the text.

Explanation Models

As mentioned above, regression analyses achieve two primary goals: explanation and estimation. The goal of explanation is to examine how much variance in the dependent variable is explained by the included independent variables. To examine this, we report two statistics: the R^2 and the R^2 change. The R^2 statistic is used in regression analyses to express the amount of variance explained in the dependent variable by the included independent variables. So, the R^2 can be interpreted as the percentage of the dependent variable explained by the included independent variables. Following the same logic, the statistic of R^2 change shows the additional percent explained with each new additional set of independent variables included.

Table B2. Percentage of Variance Explained in Full Model

Variable	R^2 Change	R^2
Offense	25.6***	25.6
OV Level	19.6***	45.2
PRV Level	23.1***	68.2
Habitual Status	2.6***	70.9
Departure	17.9***	88.8
Conviction Method	0.5***	89.3
County	0.1***	89.6
Demographics	0.1***	89.6

Table B2 displays the findings regarding the percent of variance explained in the full model, which can help us understand the relative importance of variable groups as well as how much of the variance in sentence length is explained. In the full model, what offense was committed explains the most variance—more than 25% of the differences observed in sentence length. This is followed by the two other variable groups we would anticipate have substantial explanatory power: the offense severity (OV level) and prior criminal history (PRV level). These elements explain an additional 19.6%, and 23.1%, respectively. Comparatively, habitual status contributes only 2.6% of the variance explained.[5] Taking these four variable groups together we can explain more than 70% of the variance in sentence length, with about 30% unexplained by guidelines factors.

In a perfect world, just the guidelines factors would explain 100% of the variance. In reality, other variables also affect sentence length to a greater or lesser extent, including the fact that judges often reject the guidelines recommendation and choose to depart from it. The addition of departures below and above the guidelines range explains an additional 18 percent of the variance, bringing the cumulative total to 88.8%.

From a research perspective, explaining nearly 90% of the outcome being examined is considered a great success. However, it must be remembered that the sentencing guidelines were constructed to incorporate the factors that judges have always considered most important in sentencing: the nature of the crime, how the defendant committed it, and the defendant's prior record. These three factors collectively explained only 68.2% of the variance. The defendant's habitual status reflects a prosecutor's choice to

RECEIVED by MSC 4/12/2022 1:07:49 PM

place additional emphasis on the prior record. Departures reflect a judge's choice not to follow the guidelines recommendations.

The remaining three variable groups, including the method of conviction, county of conviction and defendant demographics, cumulatively explain less than one additional percent, suggesting that when looking at the variance in sentence length, these elements do not comprise a substantial impact—which does not mean they are unimportant. In individual cases, these elements can have substantial impact on sentence length. We will explore this possibility in the next section, which examines the regression coefficients that can be used to estimate sentence length. Overall, the variables included in our regression analyses explain 89.6% of the variance in sentence length for life-maximum offenses, meaning that there is still 10.4% of the variance that left unexplained.

The additional explanation model tables, B3, which examines the percent of variance explained in each offense separately, and B4, which examines each county separately, can be interpreted in the same way as above. Please note, in each of these analyses, we are examining subgroups of our data, so the population parameters change, as do some of the variables. In the offense-specific models, offense is not included as a variable as it is already accounted for. In the county-specific models, offense is included as a variable group, and county is not for the same reason.

Breaking down the regression analyses by offense type and county allows us to look for differences within these groups. Specifically, it allows us to examine specific differences that may be overshadowed in the full model because so many of the sentences are for armed robbery and from Wayne County. For example, in looking at the offense-specific analyses, we are able to compare the relative importance of PRV level and OV level to find that they collectively explain only 25% of the variance in sentence length for murder sentences but 65% of the variance in robbery sentences. Similarly, examining county differences the guidelines factors explain only 66% of the difference in sentence length in Wayne County, but approximately 85% of the variance in Berrien, Calhoun, Kalamazoo, Muskegon, and Washtenaw counties.

Table B3. Percentage of Variance Explained by Offense Type								
	M2		AWIM		CSC		RA	
Variable	R^2 Change	R^2	R^2 Change	R^2	R^2 Change	R^2	R^2 Change	R^2
OV Level	5.4***	5.4	24.5***	24.5	27.4***	27.4	29.3***	29.3
PRV Level	19.2***	24.7	23.8***	48.3	29.2***	56.6	35.8***	65.2
Habitual Status	4.8***	29.5	5.3***	53.6	3.9***	60.5	2.8***	67.9
Departure	50.5***	80	30.2***	83.8	25.9***	86.4	19.5***	87.5
Conviction Method	0.8***	80.8	1.2***	84.9	0.6***	87	0.7***	88.1
County	0.2	81	0.4**	85.3	0.8***	87.8	0.4***	88.6
Demographics	0.1	81.2	0.1	85.5	0.1	87.9	<0.1**	88.6

Table B4. Percentage of Variance Explained by County										
	Berrien		Calhoun		Genesee		Ingham		Jackson	
Variable	R^2 Change	R^2	R^2 Change	R^2	R^2 Change	R^2	R^2 Change	R^2	R^2 Change	R^2
Offense	20.4***	20.4	10.1***	10.1	19.8***	19.8	26.9***	26.9	20.8***	20.8
OV Level	34.9***	55.3	38.9***	49	22***	41.8	12.2***	39.1	27.7***	48.5
PRV Level	27.9***	83.2	32.6***	81.6	22.2***	64.1	31.4***	70.5	28.4***	76.9
Habitual Status	3.7***	86.9	3.4***	85.1	1.9***	66	1.8***	72.3	1	78
Departure	6.4***	93.4	5.5***	90.5	21.9***	87.9	17.5***	89.8	12.7***	90.7
Conviction Method	0	93.4	2.3***	92.8	0.8***	88.7	0.3*	90.2	0.6*	91.3
Demographics	0.5	93.9	.03	93.1	0.1	88.8	0.2	90.3	0.8	92.1

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table B4. Percentage of Variance Explained by County (cont'd)										
	Kalamazoo		Kent		Macomb		Muskegon		Oakland	
Variable	R^2 Change	R^2	R^2 Change	R^2	R^2 Change	R^2	R^2 Change	R^2	R^2 Change	R^2
Offense	28.5***	28.5	21***	21	18.4***	18.4	15***	15	16***	16
OV Level	27.8***	56.3	20.5***	41.5	27.1***	45.6	32.4***	47.4	36.6***	52.6
PRV Level	25.9***	82.2	29.2***	70.7	29.7***	75.3	35.7***	83.1	29.4***	82
Habitual Status	2.3***	84.5	4.2***	75	0.7***	76	1.6***	84.7	0.6***	82.6
Departure	9.7***	94.2	15.3***	90.3	13.8***	89.8	6.2***	90.9	7.6***	90.2
Conviction Method	0.4***	94.6	0.2***	90.5	1***	90.8	0.9***	91.7	1.4***	91.6
Demographics	0.3	94.8	0.1	90.6	0.4***	91.2	0.3	92	<0.1	91.6

Table B4. Percentage of Variance Explained by County (cont'd)						
	Saginaw		Washtenaw		Wayne	
Variable	R^2 Change	R^2	R^2 Change	R^2	R^2 Change	R^2
Offense	22.8***	22.8	26.7***	26.7	34.4***	34.4
OV Level	20.1***	42.8	32.9***	59.6	11.9***	46.3
PRV Level	27.3***	70.1	24.1	83.7	17.6***	63.9
Habitual Status	0.6***	70.7	0.9***	84.6	1.9***	65.8
Departure	14.1***	84.9	8.3***	92.9	23***	88.8
Conviction Method	2.5***	87.3	0.6***	93.5	0.3***	89.1
Demographics	0.4	87.8	0.4*	93.6	<0.1***	89.1

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

Estimation Models

To achieve the second goal of estimation, a different set of statistics is utilized. To aid in understanding only the interpretable result is reported.[1] Further, only statistically significant variables in the regression analysis are included. If this column is blank, the variable was not significant at the $P < .05$ level or above. The baseline sentence is the median sentence for a non-white male aged 20–29 who pleaded guilty to armed robbery in Wayne County, had no habitual enhancement or departure and scored at OV 1 and PRV level A. Each significant variable can be interpreted as an increase or decrease of the baseline sentence.[2] For example, if the RA was scored at an OV level II and PRV level C, which are the most common, we estimate that the sentence would increase 35% for the OV level, and 73% for the PRV level.

Estimated Sentence = 24 months +(24 months x 35%) + (24 months x 73%)
Estimated Sentence = 24 months + 8.4 months +17.25 months
Estimated Sentence = 49.7 months

Figure 5d in Chapter Eight shows the actual median sentence for the C-II cell for armed robbery is 51 months, demonstrating that the process of estimation can be quite accurate. The same process can be applied for the rest of the significant variables. If the person receiving this sentence had a second habitual offender enhancement, the sentence would increase by 10% of the baseline sentence and be added to the above total. If the sentence was a downward departure, we would subtract 44% from the baseline, and add 77% if it was an upward departure.

The impact of the independent variable groups that did not appear to contribute to sentence length are more pronounced here. For example, in the full model, if the conviction method was a bench or jury trial, we would estimate that the sentence would be 6% or 18% longer, respectively, than if it was settled by plea. Similarly, county did not account for much of the variance explained, but if the sentence was from Oakland County, we would estimate it to be 7% shorter than if it was from Wayne and if it was from Muskegon we would estimate it to be 7% longer. Similar to the explanation models, the defendant characteristics generally do not have a significant impact on sentence length. In the full model if a defendant is older than 29, we estimate a slight increase in sentence length. There is no significant impact of sex in the full model. The result for race estimates that white defendants would have sentences that are 2% longer than non-white defendants even though no significant differences appear by race for any specific offense.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Table B5. Interpretation for Full and Offense Regression Models[8]					
	Full	M2	AWIM	CSC	RA
Baseline Sentence	24 months	132 months	28 months	28 months	24 months
M2	228%	-	-	-	-
MA	2%	-	-	-	-
CSC	7%	-	-	-	-
OV Level II	35%	22%	-	50%	35%
OV Level III	78%	49%	46%	107%	82%
OV Level IV	125%	-	90%	163%	121%
OV Level V	177%	-	135%	239%	162%
OV Level VI	226%	-	182%	277%	219%
PRV Level B	31%	11%	24%	30%	39%
PRV Level C	73%	32%	60%	71%	91%
PRV Level D	123%	53%	88%	109%	157%
PRV Level E	173%	85%	132%	158%	216%
PRV Level F	228%	111%	177%	213%	282%
habitual2	10%	9%	15%	17%	7%
habitual3	19%	13%	17%	29%	18%
habitual4	25%	33%	26%	34%	21%
Departure Below	-44%	-45%	-44%	-46%	-43%
Departure Above	76%	59%	65%	82%	79%
Court	6%	8%	10%	-	6%
Jury	18%	12%	21%	14%	20%
Berrien	10%	-	-	19%	9%
Calhoun	8%	-	-	13%	7%
Genesee	7%	-	8%	24%	7%
Ingham	-	-	12%	10%	-4%
Jackson	7%	-	-	21%	-
Kalamazoo	-	-	-	-	-
Kent	4%	6%	-	5%	4%
Macomb	-4%	-	-7%	7%	-5%
Muskegon	7%	-	-	10%	10%
Oakland	-7%	-	-	-	-8%
Saginaw	3%	-	-	-	-
Washtenaw	-6%	-	-	-	-7%
Under Age 17	-	-	-	-	-
Age 17	-	-	-	-10%	-
Age 18–19	-	-	-	-	-
Age 30–39	2%	-	-	-	2%
Ages 40 and older	5%	-	-	-	4%
Race	2%	-	-	-	-
Sex	-	-	-11%	-	-

RECEIVED by MSC 4/12/2022 1:07:49 PM

Note that in each of these analyses, we are examining subgroups of our data so the population parameters change, as do some of the reference categories and baseline sentences. In the section that looks at each specific offense, armed robbery is no longer the reference category; there is no reference category for offense as each analysis is examining only one offense. For the regression analysis of murder, OV Level 1 is still the reference category, but only OV levels 2 and 3 are included in the regression as murder only has 3 offense levels.

On Table B6, which displays the county-specific analyses, Wayne County is no longer the reference category, as each regression is comprised of only sentences from that county. In the county regression analyses, offense is included as an independent variable again and the reference category is Armed Robbery. In these regression analyses, the baseline sentence is the median sentence for an armed robbery, OV level 1, PRV level A, non-habitual who received a sentence within the guidelines range and pleaded guilty.

As with the explanation models, breaking down the regression analyses for the estimation models by offense type and county also allows us to investigate differences more closely. As highlighted in the text and the findings, the impact of increases in OV and PRV level vary by offense. Specifically, the increases in OV level have the largest impact on CSC sentences, while increases in PRV level are most noticeable for robbery sentences. Similarly, there are differences in the impact of the independent variables by county. For example, in the county-specific regressions, we observe that the increase in sentence length for a jury trial compared to a plea is 11% in Ingham and Kent counties, and 38% in Saginaw, while there is no statistical difference in Berrien and Jackson Counties.

Important findings from the regression analyses are included in the text where applicable.

Table B6. Interpretation for County Regression Models

Variable	Berrien	Calhoun	Genesee	Ingham	Jackson	Kazoo	Kent	Macomb	Muskegon	Oakland	Saginaw	Wash	Wayne
M2	221%	201%	230%	215%	153%	235%	217%	190%	148%	232%	212%	227%	230%
AWIM	-	-	-	28%	-	-	-	-	-	-	-	-	5%
CSC	-	-	14%	23%	-	-	-	12%	-	7%	12%	-	5%
OVL 2	31%	37%	29%	30%	41%	33%	40%	30%	53%	43%	29%	36%	28%
OVL 3	81%	96%	78%	79%	95%	78%	86%	76%	103%	98%	71%	83%	65%
OVL 4	137%	151%	141%	104%	160%	126%	125%	124%	129%	147%	116%	120%	108%
OVL 5	192%	176%	191%	142%	180%	195%	193%	172%	197%	208%	140%	198%	153%
OVL 6	255%	229%	238%	200%	200%	222%	240%	213%	244%	270%	189%	279%	192%
PRVL B	27%	-	23%	19%	57%	57%	30%	30%	38%	37%	28%	35%	31%
PRVL C	86%	53%	58%	60%	79%	103%	86%	72%	87%	82%	68%	72%	72%
PRVL D	136%	105%	98%	97%	149%	169%	138%	134%	144%	144%	115%	137%	119%
PRVL E	162%	155%	153%	153%	203%	248%	186%	212%	202%	183%	170%	164%	169%
PRVL F	250%	186%	176%	233%	191%	271%	261%	260%	299%	254%	251%	274%	213%
Hab 2	33%	14%	-	16%	35%	-	21%	-	-	-	-	-	18%
Hab 3	38%	16%	20%	-	-	24%	21%	-	29%	17%	16%	25%	19%
Hab 4	57%	52%	23%	-	41%	15%	38%	-	34%	16%	24%		27%
Dep_Below	-43%	-34%	-52%	-45%	-55%	-38%	-45%	-41%	-46%	-42%	-39%	-39%	-44%
Dep_Above	71%	76%	87%	63%	182%	87%	93%	96%	81%	65%	95%	67%	70%
Court	-	-	26%	-	-	35%	-	-	-	19%		19%	5%
Jury	-	30%	24%	11%	-	15%	11%	24%	26%	26%	38%	19%	12%
Under 17	-	-	-	-	-	-	-	-	-	-	-	-	5%
17	-11%	-	-	-	-	-13%	-6%	-9%	-	-	-	-	-
18-19	-	-	-	-	-	-	-	-	-	-	-	-	-
30-39	-	-	-	-	-	-	-	11%	-	-	-	-	2%
40+	-	11%	-	-	-	-	-	7%	-	-	-	14%	7%
Race	-	-	-	-	-	-	-	-	-	-	-	-	-
Sex	-	-	-	-	-	-	-	-	-	-	-13%	-	-
Baseline Sentence[9]	21 months	21 months	35 months	21 months	21 months	21 months	24 months	21 months	21 months	21 months	21 months	21 months	24 months

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

[1] Allison supra note 8 in Appendix A at 3.

[2] Outliers were identified using casewise diagnostics and removed for each analysis. In the full model, 198 sentences were cut. An additional 68 sentences were removed as outliers for the offense and county-specific models for a total of 266 outlier sentences removed. Examination of the excluded outliers found that:

- 36 sentences (13.5%) were for M2, 44 (16.5) for AWIM, 30 (11.3%) for CSC, and 156 (58.6%) for RA.
- The majority of these sentences resulted from guilty pleas (81.6%): 15.8% were imposed after jury trials, and 2.6% were imposed after bench trials.
- Less than 10% of the sentences excluded were within guidelines range. Almost 70% (68.4%) were downward departures: the remaining 21.8% were upward departures. Excluding in range sentences:
    - For the 36 M2 sentences, 66.7% were downward departures, and 2.8% were upward departures.
    - For the 44 AWIM sentences, 66% were downward departures and 32% were upward departures.
    - For the 30 CSC sentences 50% were downward departures and 43% were upward departures.
    - For the 156 RA sentences, 73% were downward departures and 19.2% were upward departures.
- More than half (55.3%) of the outlier sentences were from Wayne County, followed by Macomb (10.5%), Kent (9%), and Oakland (7.9%).
    - Wayne County had the highest rate of downward departures at 75%, with 2.7% in range, and the remaining 22.4% were upward departures.
    - Washtenaw County had the highest rate of upward departures at 66.7%, with the remaining 33.3% as downward departures. However it should be noted that Washtenaw County only had 9 outlier sentences that were excluded.
- Eighty-five percent of the sentences did not have a habitual enhancement, 3.4% were habitualized as a second, 3% as a third, and 8.6% as a fourth habitual offender.

[3] Ostrom supra note 14 in ch 3 at appendix 4b

[4] Mertler, Craig A., and Rachel V. Reinhart. 2005. *Advanced and Multivariate Statistical Methods: Practical Application and Interpretation.* Los Angeles: Pyrczak Publishing.

[5] This is partially due to the order in which the variables were input into the regression. Habitual status is closely correlated with PRV level. When habitual status was prioritized, the impact was greater at 14.4% and reduced the explanatory power of PRV level to 13.3%.

[1] The full regression models are available in the supplemental documentation.

[2] Ostrom supra note 14 in ch 3 at appendix 4b. The Interpretation and Advantage of using Log.

[8] As seen in figures 5b and 5c in Chapter Eight, and mentioned in note 3, Chapter Eight: the median sentence in the A-I cell for AWIM and CSC sentences were higher than the max-min. This resulted in no sentences that fit the parameters for AWIM, and only one sentence for CSC. For this reason, the midpoint of the A-I cell is used as the baseline sentence for these two offenses. Other options were explored:

- such as using the constant in a regression analysis conducted on the dependent variable of sentence length that was not logged, however this resulted in baseline sentences that were upward departures
- averaging the median sentences in the adjoining cells,
- and using the overall median for the cell (which would be the RA median).

The midpoint was selected as it was similar to the solution of averaging the median sentence of the adjoining cells and aligned with the findings that AWIM and CSC sentences start out slightly longer than RA sentences.

[9] There were no sentences that fit the baseline parameters for Berrien, Calhoun, Ingham, and Saginaw counties. The median for all counites was substituted.

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

APPENDIX C. ADDITIONAL TABLES

Table C1. Racial Demographics of Counties and Sentences

	Percent of Population			Percent of Sentences	
	White (includes Hispanics)	Non-White	Hispanic	White	Non-White
Berrien	79.7	20.3	5.8	41.0	59.0
Calhoun	81.4	18.6	5.5	55.7	44.3
Genesee	75.3	24.7	3.6	33.8	66.2
Ingham	75.6	24.4	8.0	44.6	55.4
Jackson	87.6	12.4	3.6	54.9	46.1
Kalamazoo	81.2	18.8	5.2	42.3	57.7
Kent	82.2	17.8	10.8	43.7	56.3
Macomb	80.3	19.7	2.8	45.6	54.4
Muskegon	81.2	18.8	5.8	45.4	54.6
Oakland	75.3	24.7	4.3	36.5	63.5
Saginaw	76.3	23.8	8.7	33.2	66.8
Washtenaw	74.2	25.8	4.9	36.8	63.2
Wayne	54.6	45.4	6.1	14.5	85.5

Table C2. High and Low Minimum Sentences by County and Offense

	M2		AWIM		CSC		RA	
	Low	High	Low	High	Low	High	Low	High
Berrien	12	42	5.5	75	1.5	75	1.9	40
Calhoun	8.5	70	4.3	50	1.8	62.5	1.8	50
Genesee	2.5	100	1	62.5	1.9	60	1	75
Ingham	6	50	6	75	1.4	50	1.4	33.3
Jackson	7.5	100	4.5	50	4.3	47.5	1	30
Kalamazoo	3	60	1.7	70	1.8	40	1	60
Kent	8.5	60	1.3	50	0.5	50	0.8	60
Macomb	4.8	60	1	46	1	62	0.5	62
Muskegon	7.5	75	4	46.8	2	70	1.2	56.3
Oakland	2	100	1.5	99	1	75	0.4	60
Saginaw	5	100	1.5	75	1.9	75	0.3	75
Washtenaw	6	60	1.5	30	1	56	1	56
Wayne	1	100	<0.1	100	0.8	70	0.3	75
CMIS Data								

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

APPENDIX D. HYPOTHETICAL CASE COMPARISON CHARTS

Comparison Chart 1: Second-Degree Murder, Example 1

Facts

D is 31 years old (b. 1990). His only priors are two adult misdemeanors and a juvenile adjudication for retail fraud. On May 1, 2020, V shot into D's home because he had a dispute with D's brother. The only one home was D's mother who was sitting on a couch watching TV. A bullet hit her in the leg causing a non-life-threatening injury. She called 911 herself and was taken to the hospital.

When D learned what happened, he immediately got two friends and went looking for V. D had a hunting rifle; the two friends both had handguns. They found V hanging out on a street corner with three friends of his own. D called V to come over to the car where D could confront V about the prior shooting. V refused to come over. He cursed D and his family loudly but did not display a weapon. D shot and killed V. D's friends fired shots at V's friends but no one else was injured.

D was charged with first-degree premeditated murder, three counts of assault with intent to murder and felony firearm. He pleaded guilty to second-degree murder and possession of a firearm during the commission of a felony (felony firearm).

Location	Breadth of Range (in mos)	Presumptive	Mitigated	Aggravated	Mandatory Minimum	Percent of Minimum to be Served	Statutory Maximum	Additional Firearm Penalty
Michigan[a]	120	180-300/L				100	Life	24
Kansas[b]	20	176	166	186		85	186	
Minnesota[c]	106	306 [261–367]				67	480	36
N Carolina[d]	48	192–240	144–192	240–300		100	120% of min	
Oregon[e]	n/a	Life no parole			300	100	Life	
Utah[f]	n/a	240			180	Variable	Life	
US[g]	58	235–293				85	Life	60

[a] MI: Murder II, MCL 750.317; indeterminate sentence; separate first-offense felony firearm charge has mandatory consecutive determinate 24 mos, MCL 750.227b.
[b] KS: Murder II, 21-5403; using firearm in a person felony presumes prison but allows non-prison sentence if findings are made about available treatment or the promotion of offender reformation. KSA 21-6804(h). Maximum is highest sentence in the applicable grid block.
[c] MN: Murder II, Minn Stat 609.19; mandatory minimum 36 mos up to the max for first firearm offense, Minn Stat 609.11, subd 5(a).
[d] NC: Murder II, GS 14-17(b)(1); only min range is scored, max is 120% of the minimum chosen, NC Manual at 28-29; earned credits for prison work or programs apply only to maximum and cannot reduce the minimum; consecutive sentence consideration applies for the separate firearm conviction.
[e] OR: Murder II, ORS 163.115; life with potential second look at 25y (300 mos) if D >15 at time of offense, ORS 163.115(5); mandatory min set by ORS 137.700(2)(a)(A); good time barred by ORS 421.121(1) -- excluded as second-degree murder and having a mandatory minimum.
[f] UT: Murder, 76-5-203; indeterminate sentence, number represents typical prison stay, Guidelines 17; at least 4 mos. credit for program participation unless life sentence, 77-27-5.4(3).
[g] US: Murder II, 18:1111; additional 5 year mandatory sentence for firearm used in a violent crime, 18 USC 924(c).

RECEIVED by MSC 4/12/2022 1:07:49 PM

Comparison Chart 2: Second-Degree Murder, Example 2

Facts

D is 20 years old. Since he was 15 he accumulated two juvenile adjudications and one adult conviction, all for breaking and entering an unoccupied building.

While on adult probation D committed first-degree home invasion. He believed that the 60-year-old woman who lived in the house was at work but she had taken the day off. D had managed to pocket $1,200 worth of jewelry before V heard him and came to investigate. D was not armed but when V turned to call the police, he struck her hard. V hit her head on a chest of drawers as she fell. D took the cell phone from her hand and ran. V's family found her unconscious five hours later. She died of a brain hemorrhage the next day.

D was charged with first-degree felony-murder. He pleaded guilty to second-degree murder.

Location	Breadth of Range (in mos)	Presumptive	Mitigated	Aggravated	Mandatory Minimum	Percent of Minimum to be Served	Statutory Maximum
Michigan[a]	210	315-525/L				100	Life
Kansas[b]	25	234	221	246		85	246
Minnesota[c]	127	366 [312-439]				67	480
N Carolina[d]	63	254-317	190-254	317-397		100	120% of min
Oregon[e]	n/a	Life no parole			300	100	Life
Utah[f]	n/a	252			180	Variable	Life
US[g]	52	210-262				85	Life

[a] MI: Murder II, MCL 750.317.
[b] KS: Murder II, 21-5403. The maximum is the highest sentence in the applicable grid block.
[c] MN: Murder II, Minn Stat 609.19.
[d] NC: Murder II, GS 14-17(b)(1); only min ranges scored, max is 120% of the minimum chosen, NC Manual 28-29; earned credits for prison work or programs apply only to maximum and cannot reduce the minimum.
[e] OR: Murder II, ORS 163.115; murder sentence to follow remaining probation term being served at the time of the offense; life in prison with potential second look at 25 years (300 mos) if D > 15 at time of offense, ORS 163.115(5); mandatory minimum set by ORS 137.700(2)(a)(A); good time barred by ORS 421.121(1) -- excluded as second-degree murder and having a mandatory minimum.
[f] UT: Murder, 76-5-203; indeterminate sentence, number represents typical prison stay, Guidelines 17; mandatory min 15 yrs to life, 76-5-203(3)(a); at least 4 mos. credit for program participation unless life sentence, 77-27-5.4(3);
[g] U.S.: Murder II, 18:1111.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Comparison Chart 3: Assault with Intent to Murder, Example 1

Facts

D is 37 years old. In the last five years he has had two convictions for domestic assault, one for felonious assault and two for operating while intoxicated, 3rd offense.

In the evening of the offense, D got drunk and then started railing at his wife. First he slapped her. Then he choked her until he was interrupted by their 12-year-old son. D then turned on the boy, striking and knocking down the child before fleeing the house. He was combative with the police who found and arrested him.

D was charged with assault with intent to murder, assault with intent to commit great bodily harm, resisting arrest, and being a fourth offender. He pleaded guilty to assault with intent to murder.

Location	Breadth of Range (in mos)	Presumptive	Mitigated	Aggravated	Mandatory Minimum	Percent of Minimum to be Served	Statutory Maximum
Michigan[a]	114	171–285				100	Life
Kansas[b]	26	233	221	247		85	247
Minnesota[c]	54	158 [135–189]				67	240
N Carolina[d]	63	251–314	189–251	314–393		100	120% of min
Oregon[e]	9	121–130			90	100	240
Utah[f]		108			60	Variable	Life
U.S.[g]	52	210–262				85	240

[a] MI: Assault with Intent to Murder, MCL 750.83.
[b] KS: Attempt (21-5301(c)(1)) to Commit Murder II (21-5403).
[c] MN: Assault I, Minn Stat 609.221.
[d] NC: Attempt (GS 14-2.5) to Commit M2 (GS 14-17); only min ranges scored, max is 120% of the minimum chosen, NC Manual 28-29; earned credits for prison work or programs apply only to maximum and cannot reduce the minimum.
[e] OR: Assault I, ORS 163.185; mandatory minimum set by ORS 137.700(2)(a)(G); statutory maximum set as felony Class A, ORS 161.605(1); good time barred as a charge with a mandatory minimum, ORS 421.121(1).
[f] UT: Attempt Murder, 76-5-203; 5 yrs to life, 76-3-203(1); indeterminate sentence, number represents typical prison stay, Guidelines 17; at least 4 mos. credit for program participation unless life sentence, 77-27-5.4(3).
[g] US: Assault with Intent to Murder, 18 USC 113.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Comparison Chart 4: Assault with Intent to Murder, Example 2

Facts

D is 17 years old. He has juvenile adjudications for breaking and entering a building, home invasion 2nd degree and delivery of a schedule 1 controlled substance < 50 grams. His single mother is in prison for drug delivery and he lives with his grandmother. While still on probation to the juvenile court D joined a gang that is competing for control of a neighborhood with another gang. D and two fellow gang members caught V, a member of the rival gang, selling drugs in their neighborhood. They pulled V into a vacant house where they couldn't be seen or heard. They beat V with their fists and bats and kicked him in the back, breaking his spine and leaving him paralyzed from the waist down. D said he never intended to kill anyone and as a new recruit he was just doing as ordered by older members. He went to trial and was convicted as charged of assault with intent to murder.

Location	Breadth of Range (in mos)	Presumptive	Mitigated	Aggravated	Mandatory Minimum	Percent of Minimum to be Served	Statutory Maximum
Michigan[a]	150/L	225-375/L				100	Life
Kansas[b]	10	88	82	92		85	92
Minnesota[c]	38	134 [118–156]			1yr + 1 day	67	240
N. Carolina[d]	32	125–157	94–125	157–196		100	120% of min
Oregon[e]	4	111–115			90	100	240
Utah[f]		96			60	Variable	Life
US[g]	73	292-365				85	240

[a] MI: Assault with Intent to Murder, MCL 750.83.
[b] KS: Attempt (21-5301(c)(1)) to Commit Murder II (21-5403); gang involved presumes prison but allows non-prison sentence if findings are made about available treatment or the promotion of offender reformation. KSA 21-6804(k)(1).
[c] MN: Assault I, Minn Stat 609.221; mandatory min 1 yr + 1 day for dangerous weapon other than a firearm, Minn Stat 609.11, subd 4; and +24 month enhancement for gang involvement, increased guidelines from presumed 110 months, discretionary range 94-132, spread 38, Minn Stat 609.229 subd 3(a).
[d] NC: Attempt (GS 14-2.5) to Commit M2 (GS 14-17); only min ranges scored, max is 120% of the minimum chosen, NC Manual 28-29; earned credits for prison work or programs apply only to maximum and cannot reduce the minimum.
[e] OR: Assault I, ORS 163.185, mandatory min in ORS 137.700(2)(a)(G); statutory max as felony Class A, ORS 161.605(1); good time barred as a charge with a mandatory minimum, ORS 421.121(1).
[f] UT: Attempt Murder, 76-5-203; 5 yrs to life, 76-3-203(1); indeterminate sentence, number represents typical prison stay, Guidelines 17; at least 4 mos. credit for program participation unless life sentence, 77-27-5.4(3).
[g] U.S.: Assault w/Intent to Murder, 18 USC 113.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Comparison Chart 5: Criminal Sexual Conduct 1st Degree, Example 1

Facts

D, age 34, was V's mother's live-in boyfriend. He has no prior convictions, adult or juvenile.

When V was 11, D began buying her clothes and other gifts. When V was 12, D began going into her room at night several times a month and lying next to her in bed. He told her this helped him relax so he could sleep but they should keep it a secret because it would hurt her mother's feelings. After a few months, he began stroking her body all over. After a few more months, he progressed to fondling her genital area. After several more months, D began to insert his finger into V's vagina. This went on for four months. When V turned 13, D began inserting his penis instead of his finger. After this had happened three times, V confided in the mother of a friend who called social services.

D was charged with seven counts of CSC 1 for digital penetration of a child under 13 and three counts of CSC 1 for vaginal penetration of a child 13–16 in the same household. D pleaded to four counts of CSC 1 <13 and two counts of CSC 1, 13-16.

Location	Breadth of Range (in mos)	Presumptive	Mitigated	Aggravated	Mandatory Minimum	Percent of Minimum to be Served	Statutory Maximum
Michigan[a]	84	126–210			300 (ct 1-4)	100	Life
Kansas[b]	n/a	Life	n/a	n/a	300	85	300
Minnesota[c]	35	180 [180–215]				67	360
N Carolina[d]	48	192–240	144–192	240–300		100	120% of min + 60 mos
Oregon[e]	2	58–60			100	100	240
Utah[f]		192			300	Variable	Life
US[g]	52	210–262			360	85	Life

[a] MI: CSC I, V <13 750.520b(1)(a) & CSC I, V 13-16 750.520b(1)(b); lifetime electronic monitoring 750.520b(2)(d).

[b] KS: Rape, KSA 21-5503; V's age creates off-grid felony, mand Life with mand min 25 yrs until parole eligibility, KSA 21-6627(a)(1), up to 300 mos, *id*. at (b)(2)(B); range applies if downward departure (presumptive 155, mitigated 147, aggravated 165, range 18), *id.* at (d)(1), DRM at 6-7; lifetime electronic monitoring, KSA 22-3717(u)&(v); good time barred, KSA 21-6623; crime of extreme sexual violence limits downward departure range, KSA 21-6818(a).

[b] MN: CSC I, Minn Stat 609.342; includes 25% statutory enhancement for Criminal Sexual Predatory Conduct, Minn Stat 609.3453 (base score: presumptive 144 range 144-172).

[c] NC: Rape I, GS 14-27.24; only min ranges scored, max for major sex crimes is 120% of min (round to highest month) plus 60 mos (GS 15A-1340.17(f); NC Manual at 5); earned credits for prison work or programs apply only to maximum and cannot reduce the minimum.

[d] OR: Rape I, ORS 163.375, mandatory minimum set by ORS 137.700(2)(a)(K); if V <12, mandatory minimum would be 300 months, ORS 137.700(2)(b)(D); good time barred as having a mandatory minimum, ORS 421.121(1); max set as felony class A, ORS 161.605.

[f] UT: Rape of a Child, 76-5-402.1; mand 25 yrs to life; indeterminate sentence, number represents typical prison stay, Guidelines at 17; at least 4 mos. credit for program participation unless life sentence, 77-27-5.4(3).

[g] U.S.: Aggravated Sexual Abuse with Children, 18 USC 2241(c).

RECEIVED by MSC 4/12/2022 1:07:49 PM

Comparison Chart 6: Criminal Sexual Conduct 1st Degree, Example 2

Facts

D is 29 years old. He has a juvenile adjudication for home invasion 2nd degree and adult convictions for aggravated stalking, home invasion 1st degree, and CSC 3rd degree.

D was on parole for the last conviction when he saw V at a restaurant. He followed her home and waited outside until all the lights were out. Then he broke in. V was in her bed asleep. D woke her by putting a knife to her throat. He removed her nightgown, bound her hands behind her back, and taped her mouth. He forcibly penetrated her vaginally and anally, causing bruising and tearing. Then he sat in the room watching her while she lay helpless on the bed. After 30 minutes, he took the tape from her mouth, forced her to perform oral sex with the knife to her throat, then left.

D was charged with three counts of CSC 1st and with being a fourth offender. The prosecutor made no plea offer. A jury convicted D of all counts.

Location	Breadth of Range (in mos)	Presumptive	Mitigated	Aggravated	Mandatory Minimum	Percent of Minimum to be Served	Statutory Maximum
Michigan[a]	630/L	270–900/L			300	100	Life
Kansas[b]	61	620	592	653		85	653
Minnesota[c]	63	180	-----	216		67	360
N Carolina[d]	97	386–483	290–386	Life, no parole		100	120% of min + 60 mos
Oregon[e]	9	121–130			100	100	240
Utah[f]		252			Life, no parole	100	Life, no parole
US[g]	65	262–327				85	Life

[a] MI: CSC 1, MCL 750.520b(1); lifetime electronic monitoring 750.520b(2)(d). Mandatory 25-year minimum for 4th habitual offender because sentencing offense was "serious" and at least one prior was "listed." MCL 769.12(1) (a).

[b] KS: Rape, KSA 21-5503; persistent sex offender (prior crime of extreme sexual violence) enhancement presumes prison at double the maximum presumptive term but does not apply if sentencing offense is severity level 1 or 2. KSA 21-6804(j). Crime of extreme sexual violence limits downward departure to no less than half of center range, KSA 21-6818(a).

[c] MN: CSC I, Minn Stat 609.342. For dangerous offender who commits third violent crime sentence must be at least length of presumptive. May also impose an aggravated durational departure from presumptive sentence (180) to the 30 yr max if D at least age 18 at time of current offense with no probation, parole, discharge, work release, Minn Stat 609.1095.

[d] NC: Rape I, GS 14-27.24; only min ranges scored, max for major sex crime is 120% of min (round to highest month) plus 60 mos. (GS 15A-1340.17(f); NC Manual at 5); earned credits for prison work or programs apply only to maximum and cannot reduce the minimum.

[e] OR: Rape I, ORS 163.375; mandatory min set by ORS 137.700(2)(a)(K); good time barred as having mandatory minimum, ORS 421.121(1); maximum set by felony class A, ORS 161.605.

[f] UT: Rape, 76-5-402; indeterminate sentence, 252 represents typical prison stay, Guidelines at 17; good time barred by sentence: life no parole, 77-27-5.4(3)(b); program participation credit unavailable for life sentence, 77-27-5.4(3); statute mandates life without parole for prior grievous sexual offense (CSC 3).

[g] US: Aggravated Sexual Abuse (force or threat) 18 USC 2241(a), any term of years or life.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Comparison Chart 7: Armed Robbery, Example 1

Facts

D is 25 years old. He has three prior misdemeanors and two juvenile adjudications, one for larceny from a person and one for breaking and entering an unoccupied building. At age 21, he got two adult convictions, for unarmed robbery and felony firearm, for which he received probation plus two years in prison.

D and his co-D held up a gas station late at night. Both had handguns. They took money from the cash register and left. However, the co-D realized that the clerk recognized him from high school so the co-D went back inside and shot and the clerk, killing him.

The co-D was charged with first-degree felony-murder and pleaded to second-degree murder. D was also charged with felony-murder but pleaded to armed robbery and felony firearm.

Location	Breadth of Range (in mos)	Presumptive	Mitigated	Aggravated	Mandatory Minimum	Percent of Minimum to be Served	Statutory Maximum	Additional Firearm Penalty
Michigan[a]	150	225–375/L				100	Life	60
Kansas[b]	22	216	206	228		85	228	
Minnesota[c]	23	68 [58–81]				67	240	60
N Carolina[d]	17	67–84	51–67	84–105		100	120% of min	
Oregon[e]	6	66–72			90	100	240	
Utah[f]		120			60	Variable	Life	
US[g]	23	92–115				85	240	60

[a] MI: Armed Robbery, MCL 750. 529. separate second-offense felony-firearm charge has mandatory consecutive determinate 5 yrs, MCL 750.227b.
[b] KS: Aggravated Robbery, KS 21-5420(b); firearm used in a person felony requires prison but allows non-prison sentence if findings are made about available treatment or the promotion of offender reformation. KSA 21-6804(h).
[c] MN: Aggravated Robbery I, Minn Stat 609.245, subd 1; second firearm offense mandatory minimum 5 yrs, Minn Stat 609.11, subd 5(a); aggravated durational departure as a dangerous and repeat offender, up to the statutory maximum sentence with ineligibility for probation, parole, discharge, or work release, Minn Stat 609.1095.
[d] NC: Robbery, GS 14-87; only min ranges scored, max is 120% of the minimum chosen, NC Manual at 28-29; earned credits for prison work or programs apply only to maximum and cannot reduce the minimum.
[e] OR: Robbery I, ORS 164.415, mandatory minimum set by ORS 137.700(2)(a)(R); statutory maximum set as Felony Class A, ORS 161.605(1); up to 20% good time barred by ORS 421.121(1), excluded by mandatory minimum.
[f] UT: Aggravated robbery, 76-6-302; first-degree felony, min 5 yrs, max life, 76-3-203; indeterminate sentence, number represents typical prison stay, Guidelines 17; at least 4 mos. credit for program participation unless sentenced to life, 77-27-5.4(3);
[g] U.S.: Armed Robbery, 18 USC 1951; add'l. 5 yr mandatory sentence for firearm used in a violent crime, 18 USC 924(c)

RECEIVED by MSC 4/12/2022 1:07:49 PM

Comparison Chart 8: Armed Robbery, Example 2

Facts

D is 27 years old. He is married with two young children, one of whom has a disability that creates substantial medical bills. He has no prior record and worked at the same factory for five years. However he was laid off when the factory closed and was unable to find steady employment for over six months.

D decided to commit a series of robberies. He went to liquor stores in three different counties. Each time he waited until just before closing when the store was empty. He used a realistic-looking toy gun and told each clerk to give him all the money from the till. The information about each offense is as follows:

County A: The clerk complied and gave D $1,100. D was charged with armed robbery and pleaded to unarmed robbery.
Offense – Sept 1 / Conviction – Nov 5 / Sentence – Nov 19

County B: The clerk complied and gave D $740. During the transaction a stock boy came in from the back room so D aimed the gun at him too. D was charged with two counts of armed robbery and pleaded to one count of armed robbery.
Offense – Sept 3 / Conviction – Dec 15 / Sentence – Jan 11

County C: The clerk realized that D's gun was a fake. He pulled a real weapon from under the counter, called the police and held D at gunpoint until the police arrived. D was charged with armed robbery. He pleaded guilty as charged in exchange for a promise not to add a 3rd habitual offender charge.
Offense – Sept 9 / Conviction – Jan 13 / Sentence – Jan 23

The guidelines scoring for the County B sentence is:

Location	Breadth of Range (in mos)	Presumptive	Mitigated	Aggravated	Mandatory Minimum	Percent of Minimum to be Served	Statutory Maximum
Michigan[a]	34	51–85				100	Life
Kansas[b]	11	94	89	100		85	100
Minnesota[c]	19	58 [50–69]				67	240
N Carolina[d]	17	67–84	51-67	84-105		100	120% of min
Oregon[e]	1	39–40			90	100	240
Utah[f]		72			60	Variable	Life
US[g]	11	46–57				85	240

[a] MI: Armed Robbery, MCL 750. 529.
[b] KS: Aggravated Robbery, KSA 21-5420(b).
[c] MN: Aggravated Robbery I, Minn Stat 609.245, subd 1.
[d] NC: Robbery, GS 14-87; only min ranges scored, max is 120% of the minimum chosen, NC Manual 28-29; earned credits for prison work or programs apply only to maximum and cannot reduce the minimum.

RECEIVED by MSC 4/12/2022 1:07:49 PM

[e] OR: Robbery I, ORS 164.415, mandatory min set by ORS 137.700(2)(a)(R); max set as felony class A, ORS 161.605(1); good time barred, ORS 421.121(1), excluded by mandatory minimum

[f] UT: Aggravated Robbery, 76-6-302; 1st degree felony, min 5 yrs, max life, 76-3-203; indeterminate sentence, number represents typical prison stay, Guidelines 17; at least 4 mos. credit for program participation unless sentenced to life, 77-27-5.4(3);

[g] U.S.: Armed Robbery, 18 USC 1951; toy gun does not support a separate gun charge under 18 USC 924(c).

RECEIVED by MSC 4/12/2022 1:07:49 PM

RECEIVED by MSC 4/12/2022 1:07:49 PM

APPENDIX H

CREW (Citizens for Racial Equity in Washtenaw) 2020 Final Report

RECEIVED by MSC 4/12/2022 1:07:49 PM

RACE TO JUSTICE

CITIZENS FOR RACIAL EQUITY IN WASHTENAW'S REPORT ON RACIAL DISPARITIES IN THE WASHTENAW COUNTY CRIMINAL LEGAL SYSTEM

CITIZENS FOR RACIAL EQUITY IN WASHTENAW (CREW)

AUGUST 2020

RECEIVED by MSC 4/12/2022 1:07:49 PM

TABLE OF CONTENTS

I.	PURPOSE AND SCOPE OF THE REPORT		2
	A.	COMMITTEE CHARGE	3
	B.	METHODOLOGY SUMMARY	4
II.	EXECUTIVE SUMMARY		5
	A.	FINDINGS REGARDING PROSECUTORS	5
	B.	FINDINGS REGARDING JUDICIAL DECISION-MAKING	7
	C.	RECOMMENDATIONS: PROSECUTORS AND THE COURT	10
	D.	DEEPER DIVE	16
III.	FINDINGS		16
	A.	FINDINGS: PATTERNS IN PROSECUTION	17
	B.	FINDINGS: PATTERNS IN JUDICIAL DECISION-MAKING	17
IV.	DATA HIGHLIGHTS		17
	A.	BACKGROUND	17
	B.	PATTERNS IN PROSECUTION	18
	C.	PATTERNS IN JUDICIAL DECISION-MAKING	25
V.	CREW'S INTERPRETATIONS AND RECOMMENDATIONS		35
	A.	RECOMMENDATIONS: PROSECUTION	36
	B.	RECOMMENDATIONS: JUDICIAL DECISION-MAKING	37
VI.	ADDITIONAL ANALYSIS		42
	A.	ADDITIONAL ANALYSIS USING CREW'S DATASET	42
	B.	FURTHER ANALYSIS USING OTHER SOURCES	45
VII.	ACKNOWLEDGEMENTS		47
	A.	BIOGRAPHICAL INFORMATION OF CREW	47
	B.	ACKNOWLEDGEMENTS	49
VIII.	TABLES & METHODOLOGY MEMO		50
	A.	ACRONYMS	50
	B.	LINK TO DATA TABLES	50
	C.	STATISTICIAN'S METHODOLOGY MEMO	50

RECEIVED by MSC 4/12/2022 1:07:49 PM

This is a unique moment of heightened awareness about the long-standing and harmful personal and societal impacts of racial inequities. Individuals, corporations, and organizations are urging state, county, and municipal entities to identify and finally address historical and present-day racial inequities and disparities in our society and systems of government. The call for real action has gone beyond law enforcement and reaches into our criminal legal system including prosecutors and judges. Over sixteen chief justices from the nation's state courts have supported reviews of the inequities of a system that appear to "value Black lives less than it values White lives."[1]

In an effort to contribute to the goal of a criminal legal system that ensures equal justice for all participants, Citizens for Racial Equity in Washtenaw (CREW) was formed to gather public data on the charging and sentencing of individuals in Washtenaw County and to assess whether the data reflected any racial disparities and the impact of any disparity on members of our community. CREW's work is inspired by and dovetails with Washtenaw County's *One Community: Advancing Racial Equity in Washtenaw County* initiative.

I. PURPOSE AND SCOPE OF THE REPORT

From policing to prosecution to cash bail and jury selection through sentencing and incarceration to parole and probation, there is overwhelming statistical evidence that People of Color in this country are discriminated against at every level of the criminal legal system.[2] At the county level, however, such evidence has been largely anecdotal.

CREW members came together to examine publicly available criminal case data from the Washtenaw County Circuit Court to assess whether there existed a basis for concern about disparities in the treatment of individuals, based on race, in charging and sentencing which would naturally impact the fair and just application of justice in our criminal legal system. As discussed in the methodology section below, CREW did not dig into every data point in Washtenaw County's criminal legal system which might contain disparities; we anticipate, however, that this report will prompt others to study these data points.

The public data that might provide insights into the experiences of Washtenaw County residents who come in contact with the county's criminal legal system is not readily available.[3] There is no public reporting of:

- The number, age or race of people criminally charged in Washtenaw County.
- What people are charged with or the outcomes.

---

[1] *We Are Part of the Problem They Protest*, New York Times https://www.nytimes.com/2020/06/16/opinion/state-supreme-courts-racial-justice.html

[2] *There's Overwhelming Evidence That the Criminal Justice System Is Racist. Here's the Proof.*, Washington Post https://www.washingtonpost.com/graphics/2020/opinions/systemic-racism-police-evidence-criminal-justice-system/.

[3] In Washtenaw County, the population is made up of African Americans 12.3%, American Indian and Alaska Native .4%, Asian 9.4%, Hispanic or Latino 4.9%, Native Hawaiian and Other Pacific Islander .1%, Two or More Races 3.6%, and whites 74.2%. U.S. Census Bureau Quick Facts, Washtenaw County, Michigan. https://www.census.gov/quickfacts/fact/table/washtenawcountymichigan/PST045219

RECEIVED by MSC 4/12/2022 1:07:49 PM

- What happens once someone in Washtenaw County is charged with a crime, including whether a plea is offered or the nature of the plea.
- How our elected judges sentence individuals, to whom they give probation, jail time or a prison sentence and for how long.

The Washtenaw County Circuit Court website does, however, include all criminal cases and their dispositions, albeit with limited searchability. CREW collected and examined certain categories of non-capital felony cases (FH cases) filed between 2017-2019 in Washtenaw County along with a review of all specified capital felony cases (FC cases) filed in Washtenaw County between 2013-2019[4]. In total, CREW studied 11 case categories.[5]

CREW's focus on the most common felony charges in Washtenaw County resulted in such serious racial disparities that we deemed it necessary to report our initial data review to the community. Our data and analysis show troubling racial differences in a variety of areas such as charging decisions, the use of the habitual offender designation, average convictions per case and sentencing.

At this stage, we did not draw any firm or definitive conclusions about implicit or explicit racial *bias* by any one person or entity, although there is a serious basis for concern that such biases, whether intentional or not, exist. Along those lines, CREW did not determine the source and nature of some of the disparities due to time and resource constraints. Instead, we see our work as a roadmap prompting further community discussion to address the racial disparities through specific actions outlined below, including deeper, broader study of systemic racism, individual biases, and other variables that underlie the significant racial disparities in the data we collected.

CREW is chaired by Alma Wheeler Smith and Linda Rexer. Additional members include Rev. Jerry Hatter, Dan Korobkin, Desiraé Simmons, and Rev. Joe Summers.[6]

A. Committee Charge

CREW adopted the following charge to describe the scope and purpose of its work:

To collect and examine all public data available for specified non-capital felony cases (FH cases) filed between 2017-2019 in Washtenaw County along with a review of all specified capital felony cases (FC

[4] CREW hired a professional statistician to do an analysis of the case records we documented and produce the tables (Tables Section) which form the data CREW studied. CREW collected all capital felony (FC) cases from 2005-2019 but chose to study a limited set of case categories from 2013-2019 because the judges currently hearing criminal cases in the Washtenaw County Circuit Court have been seated since 2013. We note, however, that Judge Kuhnke took over cases from her predecessor and began her own criminal docket in 2015. CREW also documented all non-capital felony (FH) cases from 2005-2019 and reviewed a limited set of case categories from 2017-2019 due to the volume of FH cases. Further analysis could review the entire span of time. Additionally, our research captured data that we did not use for this analysis but that could also be useful in a deeper study, e.g. fines levied, defendant's age, retained/appointed counsel.

[5] CREW limited our analysis to the following case categories: Assault, Resisting or Obstructing an Officer, Controlled Substance-Delivery, Controlled Substance-Possession, Suspended License, Weapons-Carrying Concealed, Weapons-Felony Firearm, Weapons-Possession by Felon, Armed Robbery, Assault with Intent to Commit Murder, Assault with Intent to Do Great Bodily Harm Less than Murder, and Homicide. A complete list of offenses captured in each category can be found in Section VIII (Tables) portion of this report.

[6] See Acknowledgments for brief biographical information about the committee members and acknowledgments of other community member who assisted.

RECEIVED by MSC 4/12/2022 1:07:49 PM

cases) filed in Washtenaw County between 2013-2019. Specifically, in the felony capital category, we examined cases involving charges of homicide, assault with intent to commit murder, assault with intent to commit great bodily harm less than murder and armed robbery. In the non-capital felony category, we examined cases involving charges related to weapons, controlled substances, obstruction/resisting arrest, and license suspension.

CREW determined what observations and questions the data posed and suggests which entities should be asked to consider those questions or help encourage further conversations and greater accountability.

The data drove CREW's observations and questions, such as highlighting where fairness appears to exist, where it may not exist or where additional data may inform those observations.

CREW may also decide to invite the broader community to participate by, among other things, launching listening sessions or an oral history project, such as a Story Corps Justice Lab, where community members can share their experiences with the justice system.

B. Methodology

As noted above, CREW did not collect and examine all sources of data regarding possible disparate treatment within the Washtenaw County criminal legal system largely because those sources are not publicly accessible. Given its primary goal of initiating informed discussions and greater accountability, CREW determined that focusing on data from some of the most common capital felony (FC) and non-capital felony (FH) cases would allow it to make pertinent observations, frame issues requiring immediate action, and pose questions in key areas that merit looking into. CREW documented the information contained in individual case records and turned it over to an experienced statistician who did a statistical analysis by race as illustrated in the Tables.[7]

CREW understands that it is possible there are multiple factors that might account for the inequities revealed by the data in the tables and has made no final findings regarding whether bias is the sole cause of the differences. We recognize that bias could be present even before the defendant is charged or reaches the courtroom including law enforcement's decision whether to arrest and the prosecutor's decision whether to press charges, add habitual offender levels, offer plea agreements and what level of bail to recommend. In addition, there are other potential factors such as varying fact patterns, prior offenses, pre-sentence investigation reports and sentencing guidelines which might impact the differences in outcomes but we have no basis to conclude that these account for all of the racial disparities in the data.

[7] The sentence lengths for each case incorporate concurrent and consecutive sentences for additional convictions. This is necessary to capture the actual type of sentence imposed as well as the total sentence imposed. For example, suppose a defendant was convicted of two different felonies and sentenced to concurrent terms. For the first conviction they were sentenced to 6 months in Jail, while the second conviction resulted in a 3-5 year prison sentence. In this example the second conviction determines the type and length of sentence for this case. Ultimately, the 6-month sentence for the first conviction would be served in prison concurrent to the 3-5 year sentence. Furthermore, the actual length of time before the defendant would be eligible for release would be 3 years, not 6 months. See Section VIII for a memorandum from CREW's statistician fully detailing CREW's methodology.

RECEIVED by MSC 4/12/2022 1:07:49 PM

With due regard for those factors, which must be acknowledged within a complex criminal legal system, CREW noted disturbing differences in outcomes for People of Color versus white defendants which warrant immediate attention. In fact, although many variables may be at play, some of the racial disparities are so stark that we have a basis for serious concern that systemic racism, individual biases, or both are damaging the integrity of the criminal legal system in Washtenaw County. Section IV of this report highlights those differences and Section VI poses questions for further inquiry.

II. EXECUTIVE SUMMARY

Reports that are full of numbers and statistics can end up looking like a tax return. As with tax returns, readers might skim the pages until they get to the bottom line. The bottom line here is that the statistics in this report are not just numbers -- they represent our neighbors, the person we sit alongside on the bus and greet on the street as we walk our dogs. They are us.

We as a nation cannot fully address racial inequities in the criminal legal system that lead to mass incarceration until we start fixing it at the local level. That is why CREW and others in the community did the tedious work of documenting and analyzing over 3,600 felony charges to see whether there are things that need fixing. The bottom line: Washtenaw County, we have some work to do.[8]

The breadth of the racial disparities CREW documented and the impact in terms of punishment on People of Color are so stark that CREW presents this report to our community with a call for immediate action by those institutions responsible for funding, administering and overseeing the criminal legal system in Washtenaw County. Every action we seek and recommendation we make is driven by the overarching goals of having a criminal legal system in Washtenaw County that is just, both in fact and in perception, and preventing future injustice based on racial disparities and unequal treatment.

A. Findings Regarding Prosecutors

Prosecutors have a unique role in the justice system not only because they have "the responsibility of a minister of justice and not simply that of an advocate" as laid out in the comment to Rule 3.8 of the ABA Model Rules of Professional Conduct[9] but also because they are often viewed as the most powerful actor in our criminal legal system[10]. When it comes to charging a person with a crime, the Washtenaw County prosecutor has broad discretion in deciding:

- Whether or not to bring charges against someone.
- What charges to bring, if any.

[8] Washtenaw County is often held up as exceptional with regard to its justice system. Our data suggests this is far from true.

[9] https://www.americanbar.org/groups/professional_responsibility/publications/model_rules_of_professional_conduct/rule_3_8_special_responsibilities_of_a_prosecutor/comment_on_rule_3_8

[10] "The office of prosecutor is regarded by many as the most powerful position in the U.S. criminal justice system. Prosecutors decide which cases to investigate, which suspects to charge, which charges to bring, and which penalties to pursue upon conviction." P. 9, *Report of The Sentencing Project to the Human Rights Committee Regarding Racial Disparities in the United States Criminal Justice System*, 2013.

RECEIVED by MSC 4/12/2022 1:07:49 PM

- Whether or not to designate a person as a "Habitual Offender."[11]
- Whether or not to offer a plea and the terms of a plea.
- Whether or not to dismiss charges.

In areas where prosecutors have discretion in deciding such things as whether to charge someone with a crime, what charges to bring and whether to designate a person as a Habitual Offender, the Washtenaw County prosecutors' decisions resulted in significant disparities between People of Color and whites across the board.

1. In all 11 case categories that CREW studied, prosecutors charged dramatically more People of Color than whites.

The differences were not small. They ranged from 22% to 1,150% more People of Color charged than whites. That is to say that, in Washtenaw County, a Person of Color is anywhere from 3 to 29 times as likely to be charged with one of the eleven case categories than a white person.

One example of the dramatic difference that CREW saw in the number of whites versus People of Color charged is in the Felony Firearm category of cases. In Michigan, a person can be charged with a violation of the statute which makes it a crime to use a firearm (legal or otherwise) in the commission of a felony. A prosecutor may choose to add a Felony Firearm charge when there are other felony charges and the felony involves a gun. By law, a conviction for Felony Firearm requires a mandatory 2-year prison sentence that is *in addition* to any sentence for any other conviction for the original charge. In other words, it automatically lengthens the sentence by 2 years. During the time frame we studied, prosecutors charged 25 People of Color with Felony Firearm while only 2 whites were charged. [12]

2. In 10 of the 11 categories, prosecutors filed more charges against People of Color, on average, than white people.

The difference in the average number of charges ranged from a 12.8% difference to 62.9% depending on the case category. The practice of routinely leveling more criminal charges against People of Color than against whites raises the questions of whether the Washtenaw County prosecutor's office is engaging in "horizontal overcharging" and "vertical overcharging." Horizontal overcharging describes instances when a prosecutor includes as many charges as possible (even ones that she/he may not be able to prove) in order to induce a person charged to accept a plea agreement and plead guilty. Vertical overcharging is the practice of charging the highest possible charge that carries the stiffest penalty with the intention that a prosecutor will use that highest charge to force a plea on the person charged. These practices are viewed as coercive techniques that may fall under the heading of "we've always done it this way" but are not necessarily in the interest of justice.

[11] If a person has previously been convicted of one or more felonies (or attempts to commit felonies), a prosecutor may choose to charge that person for a subsequent felony charge as a "habitual offender." If that person is convicted, s/he faces a much longer sentence that could lengthen his/her prison time by 25% - 100%.

[12] In asking why so many People of Color were charged as compared to whites in any of the case categories, we do not mean to suggest that more white people should be charged or that the criminal legal system should be more punitive.

RECEIVED by MSC 4/12/2022 1:07:49 PM

3. In 10 of the 11 categories that CREW studied, People of Color also received more convictions per case, on average, than whites.[13]

What accounts for People of Color consistently receiving more convictions per case than whites? Are more convictions the natural result of more charges? One of the reasons the statistics on the average number of convictions for People of Color versus whites are so disturbing is because more convictions per case makes it more likely a Person of Color will be charged as a Habitual Offender, and receive a significantly longer sentence, if she/he gets charged again with a felony. We discuss Habitual Offender enhancements further in Section IV (Data Highlights).

4. A Snapshot of the Data on Prosecutorial Decision-Making

CREW understands that it is possible there are multiple factors that might account for some of the inequities revealed by the data and has made no final findings regarding whether bias is the sole cause of the differences. We recognize that bias could be present even before a defendant is charged and did not study arrest data as part of our work. We believe that an effort like ours should be done to study racial disparities in officer-initiated contacts and arrest data.

Nevertheless, the door that the statistics opens for us shows a disturbing pattern in prosecutions in Washtenaw County. The data on charging issues illustrates the breadth and depth of a potentially problematic approach to justice in our county, in that disturbing racial disparities are evident across the board. At the very least, the data raises questions (although it does not answer them) about the exercise of prosecutorial discretion, whether there are cultural norms or implicit biases within or outside of the prosecutor's office that are contributing to unfairness, and whether and what kind of systemic reforms are necessary.

B. Findings Regarding the Judiciary

The role of a judge as a neutral and impartial arbiter holds a special place in our criminal legal system. The system hinges on that impartial arbiter avoiding impropriety and the appearance of impropriety. In other words, the perception of a judge's impropriety can be as harmful to the integrity of the legal system as actual impropriety. Impartiality and the perception of impartiality are critical in a judge's performance of her/his duties because judges have broad discretion in deciding:

- Whether to accept a plea agreement.
- Whether to sentence a person to probation, jail, or prison.[14]
- How to apply the sentencing guidelines upon a conviction.
- What minimum number of years a person must serve in prison before being eligible for release.
- What maximum number of years a person could serve in prison.
- Whether to go outside of the guidelines for sentencing since the sentencing guidelines became advisory in 2015.

[13] Since less than 5% of the cases we studied went to trial these convictions are largely the result of plea deals.

[14] Those convicted of certain offenses are required to fulfill their sentences in prison.

RECEIVED by MSC 4/12/2022 1:07:49 PM

- What weight to give mental health and substance abuse issues when sentencing.
- What weight to give to information contained in the presentence investigation report.
- What weight to give the prosecutor's sentencing recommendation.
- Whether to give weight to the victim's impact statement, if any.

1. CREW's Methodology for Analyzing the Statistics on Judges' Decision-Making

For each of the case categories, we looked at dispositions in the Circuit Court by an individual judge as well as the totality of the dispositions by the four judges who handle criminal cases, Chief Judge Carol Kuhnke, Judge Archie Brown, Judge Darlene O'Brien and Judge David Swartz. It is worth noting that our analysis is limited to the case categories we studied and is potentially impacted by the lack of public access to other data such as the rationale for departing from sentencing guidelines and the pre-sentence investigation reports used for each case. Notwithstanding those caveats, CREW looked at the following three areas in analyzing the statistics on how judges resolved and disposed of the charged offenses:

a. First, we studied each case category to see whether any individual judge displayed a significant racial disparity in her/his sentencing for the same charged offense. We found instances where there was a significant difference between the length of prison and jail sentences imposed on whites as compared to People of Color. We also found and noted disparities in the length of probation imposed, based on race.
b. Second, we assessed whether any of the sitting Washtenaw County Circuit Court judges were an outlier in terms of the average minimum/maximum prison or jail sentences or probationary term lengths he/she imposed for specific categories of charged cases.
c. Third, we looked to see whether a racial disparity occurred in terms of the type of punishment imposed by a judge upon conviction among the available options of probation only, a jail sentence or a prison sentence.

2. A Snapshot of the Data on Judicial Decision-Making

Our data offers numerous examples of one Washtenaw County Circuit Court judge or another being out of step with others on the court or showing disparities in sentencing against racial minorities in a single case category.[15] CREW highlighted 23 instances across the case categories we studied where an individual judge was an outlier as compared to other judges in terms of harsher sentencing or whose sentences showed racial disparities. We discuss each instance and the particular judge's (Brown, Kuhnke, O'Brien or Swartz) sentencing decision. We discuss those further in Section IV.

Of more concern, though, was whether we would find a pattern of disparities across all case categories by one or more judges or by the entire court. CREW found that one Washtenaw County judge's sentencing decisions accounted for 13 of the 23 issues we discuss below which appears to reflect a

[15] Judicial discretion is valued and should not be limited artificially; indeed, when it takes into account appropriate factors, it can enhance justice. It just should not be used a discriminatory manner.

RECEIVED by MSC 4/12/2022 1:07:49 PM

troubling pattern of disparities and/or harsher sentencing practices against racial minorities across the case categories but that the Washtenaw County Circuit Court, as a whole, did not demonstrate a pattern of racial disparity in its sentencing of our community members across all case categories we studied.

3. The summary table below provides an overview of those case categories where we found at least one instance of racial disparity or harsher sentencing practices among the judges. The table shows where instances of racial disparity or harsher sentencing practices were indicated in every case category and for which judge. Each instance is discussed further in Section IV.

Offense	Judge Brown	Judge Kuhnke	Judge O'Brien	Judge Swartz
ARO	*Disparity/Place			*Disparity/Sen *Disparity/Place
DEL	*Disparity/Sen *Disparity/Prob *Disparity/Place			*Disparity/Place
POS	*Disparity/Sen			
WCC	*Disparity/Sen *Disparity/Place	*Disparity/Prob		*Disparity/Place
WPF	*Disparity/Sen *Disparity/Prob *Disparity/Place	*Outlier		
ARM	*Disparity/Sen		*Disparity/Sen	*Disparity/Sen *Outlier
GBH	*Disparity/Sen			*Disparity/Sen
AWIM	*Outlier			

Key: Racial Disparity in Sentencing = Disparity/Sen
Racial Disparity in Placement = Disparity/Place
Racial Disparity in Length of Probation = Disparity/Prob
Outlier/Higher Minimum and Maximum Sentence = Outlier

RECEIVED by MSC 4/12/2022 1:07:49 PM

C. Recommendations: Prosecutors and the Court

The data CREW compiled sheds light on deep racial disparities within Washtenaw County's criminal legal system that members of the public have undoubtedly experienced but have never seen in documented form. In Section V, CREW sets forth recommendations for action and mechanisms to address racial inequities and discriminatory treatment in our criminal legal system because we believe the institutions central to that system must not only operate in a non-discriminatory manner in fact but must also be perceived to operate in that way. Restoring public confidence in the criminal legal system drives the recommendations set forth below. Additionally, Section VI lays out other areas of review and analysis that should be studied with the goals of better defining and understanding the extent and cause of the disparities and, importantly, of rectifying them.

To address the deeper issues in the Prosecutor's Office, we recommend:

1. Rigorous financial and administrative oversight by the Washtenaw County Board of Commissioners, including the establishment of a citizen's race equity commission chosen with input of community members.
2. The Board of Commissioners, in collaboration with the prosecutor-elect and the citizen's race equity commission, engage a neutral, third-party evaluator to study prosecutors' files, policies and procedures and make recommendations for data driven, evidence-based improvements to address racial disparities and determine whether the tools and practices employed by the Washtenaw County prosecutor's office are applied in way that is not racially discriminatory and does reflects the fair administration of justice. Issues to review and/or revise include, among other things:
    a. The disparate impact that the Habitual Offender designation has on People of Color.
    b. The charging of Weapons Felony Firearm. Are there written guidelines; if so, are they applied without regard to race?
    c. The factors used to offer plea deals and the creation of a transparent process for plea negotiations.
    d. The impact that the race of the victim might have on such things as charging decisions, conviction rates, and sentencing decisions.
    e. The use of vertical and horizontal charging practices.
    f. The factors/circumstances involved in deciding to dismiss charges including the cases where all charges are dismissed
    g. Current use of diversion (i.e., before and after filing charges) and restorative justice; the implementation of evidence-based diversion, deflection and/or restorative justice practices that are shown to improve outcomes for people involved in the County's criminal legal system in a racially neutral manner.
    h. The current performance review process for assistant prosecutors and other staff to ensure it is consistent with the racial equity considerations outlined in this report and supports the "minister of justice" role that prosecutors should play in the legal system.

RECEIVED by MSC 4/12/2022 1:07:49 PM

i. All assault arrests to identify how decisions are made regarding which charges to file (i.e., misdemeanor, felony or no charges) and whether People of Color are treated the same as whites in terms of the severity of the assault charges filed.
j. Community members' experiences in the criminal legal system.

3. The creation of an online dashboard to display data points like those documented by CREW so the public can assess whether the decisions by the Washtenaw prosecutor's office have a disparate impact on racial minorities. Citizens elect the prosecutor but have almost no data upon which to assess the performance of the office; a publicly available dashboard would provide voters and others with such data and improve the transparency of justice. A dashboard would also assist the prosecutor's office because they do not maintain aggregated data on data points such as race, age, and income.[16]

4. Developing written policies and procedures for charging decisions and sentencing recommendations made within the Washtenaw County prosecutor's office. Any such policy/procedure should be accompanied by a racial impact statement (i.e. an analysis assessing the possible racial consequences of any proposed policy before adopting it to avoid any unintended disparate racial effects.)[17]

5. Developing and implementing transparent mechanisms, processes and/or rules for exercising discretion, including a robust process to review assistant prosecutors' decisions regarding cases to ensure they are not engaging in racially weighted decisions that have a disparate impact on racial minorities.

6. Implementing a mandatory, ongoing training program on implicit and explicit racial bias for all county employees within the prosecutor's office and those individuals and entities used to investigate cases.

7. Developing and implementing regular equity audits to ensure that county employees within the prosecutor's office and others used by that office to investigate cases are not engaging in practices that have a disparate impact on racial minorities in Washtenaw County.

8. The Michigan Legislature use CREW's data on the discriminatory impact of the Habitual Offender law as well as available research to eliminate the Habitual Offender statute.

To address the deeper issues in the Court, we recommend:

1. Since state courts across the country have acknowledged that the judicial system can be part of the problem in the unequal treatment of minorities, particularly with regard to Black people in the criminal legal system, and that such discriminatory treatment contributes to mass incarceration of People of Color, we are compelled to ask why hasn't our court developed any method of

[16] In answer to a FOIA request asking whether the Washtenaw County prosecutor's office has aggregated data on such things as age and race, First Assistant Prosecutor Eric Gutenberg wrote, "This office does not have aggregated data on data points such as race or age of those charged." Email dated August 17, 2020 from Eric Gutenberg to MaryAnn Sarosi.

[17] In fact, any policy or procedure adopted regarding prosecutorial functions should be accompanied by a racial impact statement as states such as Iowa, Connecticut, Oregon, and New Jersey have done. *Report to the United Nations on Racial Disparities in the U.S. Criminal Justice System*, The Sentencing Project, 2018 https://www.sentencingproject.org/publications/un-report-on-racial-disparities/

RECEIVED by MSC 4/12/2022 1:07:49 PM

transparency in reporting these problems to the public and initiated oversight and change? Since our case data came from the court's own public website, but in a format that takes hundreds of hours to collect and analyze, it's natural for us to wonder why the court hasn't performed an equity audit in the past; or, if it has, why such an audit hasn't been made public so voters can make an informed decision when they elect a judge. While the disparities revealed in our analysis pre-date the tenure of the current chief judge of the Washtenaw County Court, we strongly encourage the use of our report to assist the current court to undertake serious and transparent reforms.

2. The Washtenaw County Board of Commissioners also has an integral role to play in the reforms needed to address racially discriminatory practices in the county's criminal legal system. Not only does the Board have an interest in ensuring that the county's criminal legal system aligns with the *One Community: Advancing Racial Equity in Washtenaw County* initiative, it has oversight of the county's budget, 20% of which is allocated to judicial functions.[18] Given the responsibility of the Board as well as the Washtenaw County Court to ensure equal justice in the criminal legal system, they could jointly establish a citizens race equity commission made up of community members appointed by relevant stakeholders to:

    a. Annually compile, study, and publish statistics such as those analyzed by CREW to detect racial disparities in the Washtenaw County Courts. The data could be posted on an online dashboard described below.

    b. Address the issues and questions raised in Section VI of CREW's report. As noted in that section, some of the questions can be answered using the capital felony (FC) and non-capital felony (FH) dataset that CREW has already compiled while other questions require additional information that was not available to CREW. For example, the citizens race equity commission could examine the cases involving life sentences to understand why life sentences are given to People of Color at such a disproportionate rate compared to whites.

3. CREW noted 23 instances where a Washtenaw County judge's sentencing decision exhibited racial disparity or was an outlier in terms of handing down harsher sentences than others on the court. More than half of the concerns CREW raised in this report (i.e., 13 of the 23) come from the judicial decision-making of one judge. The other 3 judges combined comprised less than half of our concerns. With those 23 instances in mind, we urge Washtenaw County Court to, among other things:

    a. Study the 23 issues further by gathering additional information on the cases involved to understand and address the source of the disparities. The study could be done under the auspices of the citizens race equity commission described above and would address the questions raised in Section VI (Additional Analysis) below.

[18] Washtenaw County 2019-2022 Preliminary Budget Summary, D-4 https://www.washtenaw.org/DocumentCenter/View/11250/2019-2022-Preliminary-Budget-Summary

RECEIVED by MSC 4/12/2022 1:07:49 PM

b. Study whether there is an imbalance among the judges in terms of defendants designated as Habitual Offenders, average number of convictions, percentage of dismissed charges, etc.

4. The disparities exhibited in cases presided over by Judge Brown are so strikingly consistent among all eight categories we highlighted and affect so many people that the court should consider taking immediate action (in collaboration with the impartial citizen's race equity commission). Therefore, we:

    a. Suggest that the Washtenaw County Court engage a neutral, outside, third-party to determine whether the racial disparities evidenced in Judge Brown's sentencing decisions are rooted in any personal bias or systemic bias. To fully capture and understand the source of the disparities, the third-party entity should be permitted and encouraged to: thoroughly review the cases filed in the case category where the issue is raised in our report during the time frame CREW studied; notify the defendants in the cases under review of the existence of CREW's report as well as the review underway by the third-party entity; solicit confidential interviews with party participants in the cases and public comment from those who have been involved and impacted; and inquire whether additional information is needed to assess the disparities in question.

    b. Believe the review should be completed in a timely manner in the interest of justice.

    c. Understand that the public could rightfully ask for assurance that the pattern of disparities in Judge Brown's sentencing decisions do not continue during the review period and believe the court has a variety of tools it could implement to instill confidence in the justice system. For example, the Washtenaw County Court could supervise/monitor his criminal cases, place a moratorium on sentencing in Judge Brown's pending criminal cases or shift his criminal caseload temporarily pending outcome of the review. Taking any of those steps would signal to the public that the court understands that judges must be neutral in fact and be publicly perceived "as an impartial dispenser of justice."[19]

5. The Washtenaw County Court could also:

    a. Issue a Local Administrative Order addressing racial bias and prejudice in conduct and judicial decisions such as sentencing.
    b. Institute a mechanism for receipt of public comment and complaints related to bias and prejudice in judicial conduct (including judicial staff) and judgments with a mechanism for investigation, review, and implementation of any remedies.

[19] *People v. Killebrew*, 416 Mich. 189, 202; 330 N.W.2d 834 (1982). Such a review would not preclude a person from filing a complaint with the Judicial Tenure Commission or taking similar action.

RECEIVED by MSC 4/12/2022 1:07:49 PM

c. Develop and implement a publicly available dashboard that is searchable and updated regularly for reporting such things as: all charges, all sentences, pleas with details of offense, and race broken down by judge. The creation of an online dashboard by the Washtenaw County Court and/or the State Court Administrative Office will assist the public in assessing whether a judge's decisions have a disparate impact on racial minorities.
d. Establish a protocol whereby all policy and procedures adopted by the Washtenaw County court regarding judicial functions in the criminal legal system should be accompanied by a racial impact assessment.
e. Provide ongoing equity/implicit bias training to all county employees at the courthouse and ensure that their performance reviews reflect any concerns about unequal treatment of courthouse users. We recognize that judges are not the only tax-funded court employees whose decisions could result in disparate treatment and outcomes. Judicial attorneys, for example, participate in decision-making and draft rulings in cases and can, therefore, impact proceedings. As with judges, they should not administer justice in a preferential manner whether in fact or by perception.[20] When a court employee's actions suggest unequal treatment of a person/persons using the court, action should be taken so Washtenaw County residents can have confidence that everyone receives equal justice under law.

6. Oversight of Racial Disparities in Judicial Decision-Making

Judges are human and not necessarily immune from the myths and biases, implicit as well as explicit, that have informed decisions and impacted our criminal legal system laws for decades, resulting in racial disparities in arrests, treatment, and incarceration of People of Color at astonishingly high rates.[21] As CREW studied the data on patterns of racial disparities related to sentencing decisions by Washtenaw County judges, we looked to see whether there was an oversight mechanisms that should have caught the disparities and possible discriminatory treatment evident in our data.

First, we looked at the court of appeals. While there is appellate review of judicial sentencing and decisions, CREW's understanding is that the court of appeals limits its review to whether there is abuse of discretion or error in applying the law to the facts of an individual case. There appears to be no effective mechanism for appellate review of overall patterns of racial disparities that could indicate discriminatory or racially biased rulings.

Second, we looked to the ethical canons embodied in the Code of Judicial Conduct. While the Michigan Supreme Court authorizes the Canons, adherence to the Canons is supervised by the Judicial

[20] A judicial attorney to one of the Washtenaw County judges suggested in a news article that she may not treat people equally when she said, "I really do keep a list of attorneys that annoy me. (There are currently 9 people on the list and you know who YOU are.)" While her comment might have been intended as a joke, it might not be perceived that way by the public and serves to undermine confidence in the justice system. http://www.legalnews.com/washtenaw/1266844

[21] Sara Beale, *The News Media's Influence on Criminal Justice Policy: How Market-Driven News Promotes Punitiveness,* 48 Wm. & Mary L.Rev. 397(2009); John J. Dilulio, Jr., *My Black Crime Problem and Ours: Why Are So Many Blacks in Prison? Is the Criminal Justice System Racist? The Answer is Disquieting,* City Journal, Spring 1996*; Hon. Harold Hood, The Race/Ethnic Bias Task Force Four Years Later—Looking Back,* 73 Mich B.J.267 (1994)

RECEIVED by MSC 4/12/2022 1:07:49 PM

Tenure Commission (JTC) which is responsible for "investigating complaints of judicial misconduct and judicial incapacity, and for recommending discipline of judges by the Michigan Supreme Court."[22]

The JTC, however, is limited in its review and investigation to conduct that violates the Code of Judicial Conduct, and Michigan's Code does not explicitly speak to racial bias or prejudice in decision-making, including that which may result in racial disparities. The American Bar Association's Model Code of Judicial Conduct, by contrast, includes a Canon explicitly prohibiting racial and gender prejudice in performance of a judge's duties. Other states have adopted such a Canon for judges, such as Canon 3 B (5) of the Virginia Canons of Judicial Conduct, which states that:

"A judge shall perform judicial duties without bias or prejudice. A judge shall not, in the performance of judicial duties, by words or conduct manifest bias or prejudice, including but not limited to bias or prejudice based upon race, sex, religion, national origin, disability, age, sexual orientation or socioeconomic status, and shall not permit staff, court officials and others subject to the judge's direction and control to do so.[23]

Given that Michigan's Canons do not explicitly cover bias and prejudice[24], we are concerned that racial disparities in sentencing People of Color may not receive the attention it deserves by the JTC. In fact, the JTC expressly states that claims involving a judge's discretionary handling of judicial duties do not fall under misconduct and are, therefore, not within the JTC's authority.[25] That leaves open the question, "Where does oversight lie for a judge's discretionary handling of judicial duties, when such acts include racial disparities in sentencing?"

Third, we considered if the public has oversight capabilities to assess whether a judge's decision-making results in racial disparities to the disadvantage of People of Color. In Washtenaw County, and indeed across Michigan, the public lacks effective mechanisms for review of a judge's record as there are no audits, scorecards or dashboards to inform the public of disparities and other patterns that emerge from the decisions of a specific judge.

Without such a mechanism, there is no basis for the public to make fully informed election decisions regarding incumbent judges, when racial disparities in decision-making or signs of bias or prejudice are important to voters. So, while the voters in Michigan elect judges, the public is largely flying blind.

---

[22] The authority and composition of the membership of the Judicial Tenure Commission comes from Article 6 Sec 30 of the Michigan Constitution adopted in 1968. Does the composition of the JTC which has been in place since 1968 (where 5 of the 9 members must be judges, 2 more are lawyers and only the remaining 2 are non-lawyers) adequately reflect input from the public or serve as an effective oversight body such that the public can have confidence in its fairness?

[23] http://www.courts.state.va.us/courts/scv/canons_of_judicial_conduct.pdf

[24] Michigan Canon 3(A)(14) does provide: "Without regard to a person's race, gender, or other protected personal characteristic, a judge should treat every person fairly, with courtesy and respect." However, the Canon does not explicitly address bias or prejudice in decision-making, including that which may result in racial disparities.

[25] State of Michigan Judicial Tenure Commission Annual Report 2019 at Page 7
http://cms4.revize.com/revize/mjtc/annual_report/docs/2019%20Annual%20Report.pdf

RECEIVED by MSC 4/12/2022 1:07:49 PM

Thus, CREW was left with the same question with which we started: How should judges in Michigan be held accountable for patterns of discriminatory judgements and/or biased treatment of individuals who come to the court?

To answer this question, we urge the Michigan Supreme Court to:

a. Examine the ways in which the Michigan Constitution protects against racial bias, discrimination, and prejudice and how those protections are evidenced in the criminal legal system.

b. Issue for public comment, a Canon that is consistent with the ABA Model Judicial Canons, which explicitly addresses the mandate that a judge's conduct (and that of her/his judicial staff) and judgments be without bias and prejudice based on all protected categories and status.

c. Provide a basis for review and audit of racial disparities in criminal sentencing and authorize the JTC to investigate such disparities for violation of the new Canon.

7. We urge the Michigan Judicial Tenure Commission to provide a mechanism for public complaints on racial disparity and bias in judicial conduct and judgements and make such data public.

D. Deeper Dive

When CREW began this effort months ago, we had no idea where the data would take us. We thought our efforts would provide a springboard into the deep waters of racial disparities in Washtenaw County's criminal legal system, but we had no idea there would be so many underwater caves to explore. In Section VI, CREW poses over thirty questions that explore those caves; resolving the questions will be necessary if the Washtenaw County criminal legal system is committed to eliminating and preventing racial discrimination.

III. FINDINGS

The public data that CREW studied reveals significant racial disparities in both the charging and sentencing of members of our community. The breadth of these disparities and the impact in terms of punishment on People of Color are so stark such that CREW presents this report to our community with a call for further informed inquiry and discussion by others in our communities and immediate action by those institutions responsible for funding, administering and oversight of the criminal legal system in Washtenaw County.

RECEIVED by MSC 4/12/2022 1:07:49 PM

A. Findings: Patterns in Prosecution

1. The data related to charges filed by the Washtenaw County prosecutor's office present disturbing patterns of racial disparities across virtually every case category that CREW examined.
2. In all 11 case categories CREW examined, prosecutors charged many more People of Color than whites (up to 29 times more).
3. In 10 of the 11 categories, prosecutors filed more charges against People of Color, on average, than whites. The difference in the average number of charges ranged from a 12.8% difference to 62.9%.
4. In 10 of the 11 categories that CREW studied, People of Color also received more convictions per case, on average, than whites.
5. In all 11 categories, prosecutors charged People of Color as Habitual Offenders at a higher rate than whites.

B. Findings: Patterns in Judicial Decision-Making

1. While we found racial disparities among individual judges, CREW did not find that the Washtenaw County Circuit Court, as a whole, demonstrated a pattern of racial disparity in its sentencing of our community members across all eight case categories we studied. That is not to say, however, we found no instances of disparity across the court. For example:
   a. Our data revealed that 44.1% of People of Color convicted of homicide were sentenced to life in prison compared to only 27.3% of white defendants convicted of homicide.
   b. Most of the judges contributed to a wide racial disparity in the average minimum/maximum sentence in armed robbery cases.
2. In addition, we highlighted 23 instances across the case categories that we studied where an individual judge was out of step with others on the court in terms of harsher sentencing or whose sentences showed racial disparities. Those instances are discussed below.
3. Over half of the instances we raise in the judicial decision-making section below relate to *one* judge (13 of the 23 issues). The other 3 judges accounted for less than half of the concerns we pointed out. The data revealed that the one judge's sentencing decisions reflected a pattern of racial disparities negatively impacting People of Color and/or harsher sentencing practices across all eight case categories.

IV. DATA HIGHLIGHTS

A. Background

1. To navigate the data tables , it is helpful to understand their design and organization. The statistical tables analyzing Washtenaw County Circuit Court records are divided into the 11 categories of cases reviewed and analyzed by CREW. (See list at footnote 5.) For each category, there is a teal colored table (reporting or summarizing the charges brought by prosecutors) and an orange table (reporting or summarizing the convictions and judicial sentencing resulting from those charges).

RECEIVED by MSC 4/12/2022 1:07:49 PM

2. Prosecutors have a unique role in the justice system not only because they have "the responsibility of a minister of justice and not simply that of an advocate" as laid out in the comment to Rule 3.8 (Special Responsibilities of a Prosecutor) of the ABA Model Rules of Professional Conduct[26] but also because they are often viewed as the most powerful actor in our criminal legal system[27]. When it comes to charging a person for an incident that may give rise to a claim of violation of the criminal statutes, the Washtenaw County prosecutor has broad discretion in deciding:

   - Whether or not to bring charges against someone.
   - What charges to bring, if any.
   - Whether or not to designate a person as a "Habitual Offender".[28]
   - Whether or not to offer a plea and the terms of a plea.
   - Whether or not to dismiss charges.

3. When it comes to sentencing a person who: a) admits guilt to one or more of the charges brought by a prosecutor; b) enters a no contest or guilty plea for a plea agreement; or c) is convicted by a jury or judge, Washtenaw County judges have broad discretion in deciding:

   - Whether to accept a plea agreement.
   - Whether to sentence a person to probation, jail, or prison.[29]
   - How to apply the sentencing guidelines upon a conviction.
   - What minimum number of years a person must serve in prison before being eligible for release.
   - What maximum number of years a person could serve in prison.
   - Whether to go outside of the guidelines for sentencing since the sentencing guidelines became advisory in 2015.
   - What weight to give mental health and substance abuse issues when sentencing.
   - What weight to give to information contained in the presentence investigation report.
   - What weight to give the prosecutor's sentencing recommendation.
   - Whether to give weight to the victim's impact statement, if any.

B. Patterns in Prosecution

1. The following tables reflect the Washtenaw County prosecutors' use of discretion in charging. In reviewing Table 1 of both the FC and the FH statistics (see Tables), we noticed that Black people were overrepresented as a percentage in the People of Color category. Black people represent

[26] https://www.americanbar.org/groups/professional_responsibility/publications/model_rules_of_professional_conduct/rule_3_8_special_responsibilities_of_a_prosecutor/comment_on_rule_3_8

[27] "The office of prosecutor is regarded by many as the most powerful position in the U.S. criminal justice system. Prosecutors decide which cases to investigate, which suspects to charge, which charges to bring, and which penalties to pursue upon conviction." P. 9, *Report of The Sentencing Project to the Human Rights Committee Regarding Racial Disparities in the United States Criminal Justice System*, 2013

[28] If a person has previously been convicted of one or more felonies (or attempts to commit felonies), a prosecutor may choose to sentence that person for a subsequent felony charge as a "habitual offender." If that person is convicted, s/he faces a much longer sentence that could lengthen his/her prison time by 25% - 100%.

[29] Those convicted of certain offenses are required to fulfill their sentences in prison.

RECEIVED by MSC 4/12/2022 1:07:49 PM

12.3% of the total population of Washtenaw County. While they make up about 52% of the People of Color in Washtenaw County, Black people account for 98.16% of the People of Color charged by prosecutors in serious capital felony cases (FC cases) and 99.4% of the People of Color charged in non-capital felony cases (FH cases).

2. The teal colored tables in each category (see Tables) describing the *charges* brought by the Washtenaw County prosecutor present disturbing patterns of racial disparities across virtually every case category of charges brought by prosecutors for criminal actions.[30]

   a. Cases charged:[31] In all 11 categories, prosecutors charged many more People of Color than whites, both in absolute numbers and in comparison, to their presence in the general population. You can see from the table below the difference ranged from 22% more to 1,150% more People of Color charged than whites[32] and a Person of Color is anywhere from 3 to 29 times as likely to be charged as a white person.

[30] CREW's data challenges the myth that Washtenaw County is exceptional in terms of racial equity. The data points to racially weighted decision-making by Washtenaw County prosecutors that is line with national studies that describe how discretionary decision-making by actors in the criminal legal system are used to more negatively impact People of Color than whites. *Report to the United Nations on Racial Disparities in the U.S. Criminal Justice System*, The Sentencing Project, 2018 https://www.sentencingproject.org/publications/un-report-on-racial-disparities/

[31] Of the 56 FC cases (about 12.6% of FC cases) in which all charges were dropped and a defendant, People of Color made up 76.8% of this group, whites made up 12.5% and 10.7% of the cases had no race information for the defendants. In the FH study, of the 140 cases (7.7% of all FH cases) in which the defendant was not convicted of any charge, People of Color accounted for 55.7% of this group, whites accounted for 41.4% and the race of 2.9% of the defendants was unknown. More charges being dropped against People of Color may suggest a lower threshold for arresting them than for a white person.

[32] CREW did not have access to arrest data; thus, we could not determine how/whether law enforcement contributed to the imbalance between those charged who were Persons of Color and white. CREW's report should prompt law enforcement leaders in Washtenaw County and the County Board of Commissioners to review and make public data on arrests.

RECEIVED by MSC 4/12/2022 1:07:49 PM

		Cases Charged			Based on their Population in Washtenaw County…
		PoC	White	Diff %	
FC Cases[33]	ARM	140	39	259.0%	A PoC is 8 ½ times as likely to be charged with ARM as a white person
	AWIM	41	10	310.0%	A PoC is almost 10 times as likely to be charged with AWIM as a white person
	GBH	35	17	105.9%	A PoC is almost 5 times as likely to be charged with GBH as a white person
	HOM	38	12	216.7%	A PoC is 7 ½ times as likely to be charged with HOM as a white person
FH Cases[34]	ARO	172	56	207.1%	A PoC is over 7 times as likely to be charged with ARO as a white person
	DEL	96	23	317.4%	A PoC is almost 10 times as likely to be charged with DEL as a white person
	POS	146	94	55.3%	A PoC is almost 4 times as likely to be charged with POS as a white person
	SLI	110	90	22.2%	A PoC is almost 3 times as likely to be charged with SLI as a white person
	WCC	108	37	191.9%	A PoC is about 7 times as likely to be charged with WCC as a white person
	WFF	25	2	1150.0%	A PoC is more than 29 times as likely to be charged with WFF as a white person
	WPF	67	8	737.5%	A PoC is almost 20 times as likely to be charged with WPF as a white person

i. What explains the pronounced tilt toward charging People of Color (primarily Black people since they comprise over 98% of the People of Color charged in the cases we analyzed) when they represent less than 30% of Washtenaw County's population? Is it possible that whites who were arrested for similar actions were charged with misdemeanors instead of felonies? Are these the only cases being brought to prosecutors by law enforcement? If so, why is that? If not, then what decisions do prosecutors make that lead them to charge People of Color overwhelmingly more than whites?

ii. The Felony Firearm (WFF) is an interesting category to examine. In Michigan, a person can be charged with a violation of the statute which makes it a crime to use a firearm (legal or otherwise) in the commission of a felony. Therefore, WFF is not generally a

[33] Armed Robbery (ARM), Assault with Intent to Commit Murder (AWIM), Assault with Intent to Commit Great Bodily Harm Less than Murder (GBH), Homicide (HOM)

[34] Assault, Resisting, Obstructing an Officer (ARO), Controlled Substance-Delivery (DEL), Controlled Substance-Possession (POS), Suspended License (SLI), Weapon-Carrying Concealed (WCC), Weapon-Felony Firearm (WFF), Weapon-Possession by Felon (WPF)

RECEIVED by MSC 4/12/2022 1:07:49 PM

stand-alone charge; a prosecutor may choose to add it when there are other felony charges and the felony involves a gun. By law, a conviction for WFF requires a mandatory 2-year prison sentence that is *in addition* to any sentence for any other conviction for the original charge. In other words, it automatically lengthens the sentence by 2 years.

As with all charges, prosecutors have total discretion to decide whether to charge a person with WFF. During the time frame we studied, 25 People of Color were charged with WFF while only 2 whites were charged (for those 2 white people, the WFF charges were dismissed.) Were firearms used in only 2 of the cases involving whites? A review of cases involving white defendants confirm that firearms were used in the commission of other felony cases, however, prosecutors chose not to add on a WFF charge.[35]

iii. Weapons Possession by a Felon (WPF) is a charge that can be filed when a person who has been convicted of a felony is found in possession of a firearm. As with other charges, it is in the prosecutor's discretion whether to file a WPF charge. The table above shows that 737% more People of Color were charged with WPF than whites in the time frame we studied. That is an astounding gap that begs the question "Why?" Is it the case that in the 1,014 non-capital felony (FH) and capital felony (FC) cases we reviewed there were only 8 cases of a white person with a prior felony conviction who possessed a firearm?

b. Average Number of Charges per Case: The broad discretion prosecutors have in charging people with crimes leaves open the possibility of "horizontal overcharging" and "vertical overcharging."[36] Horizontal overcharging describes instances when a prosecutor includes as many charges as possible (even ones that she/he may not be able to prove) in order to induce a person charged to accept a plea agreement and plead guilty. Vertical overcharging is the practice by prosecutors of charging the highest possible charge that carries the stiffest penalty with the intention that they will use that highest charge to force a plea on the person charged.

In 10 of the 11 categories, prosecutors filed more charges against People of Color, on average, than whites. The difference in the average number of charges ranged from a 12.8% difference in the average number of charges (GBH cases) to 62.9% (HOM cases).[37] In 8 of the 11 categories, the average number of charges for People of Color was more than 20% higher than their white counterparts.

[35] We reviewed court records and media of cases involving white defendants who used a firearm in the commission of a felony but were not charged with felony firearm.

[36] A.W. Alshuler, *The Prosecutor's Role in Plea Bargaining*, 36 U Chi L Rev 50, 85-105 (1968); A.M. Crespo, *The Hidden Law of Plea Bargaining,* 118 Colum L Rev 1303 (2018) (By inflating the substance of the charges beyond what the law, evidence and equities of the case supply, a prosecutor can achieve a potent lever for plea bargaining even when they are unlikely to satisfy the burden of proof at trial).

[37] In the homicide category, for example, prosecutors charged People of Color with an average of 6 charges while White defendants had an average of 3.7 charges.

RECEIVED by MSC 4/12/2022 1:07:49 PM

		Charges per Case (Avg.)		
		PoC	White	Diff %
FC Cases	ARM	4.6	3.8	20.5%
	AWIM	6.6	4.7	40.1%
	GBH	5.8	5.1	12.8%
	HOM	6.0	3.7	62.9%
FH Cases	ARO	4.2	4.5	-5.9%
	DEL	3.0	2.4	24.6%
	POS	3.3	2.7	20.5%
	SLI	3.6	3.1	15.1%
	WCC	3.0	2.3	30.9%
	WFF	5.8	4.5	28.0%
	WPF	4.4	3.4	29.1%

i. There may be explanations to account for some of the differences but the pattern across the board and the size of the differences seem to reveal a more disturbing problem. Had there been a large disparity in only one or two categories, we might have investigated those categories further to determine if there was an extreme outlier case that skewed the averages. But the disparities across the spectrum of cases leads us to believe that the gaps between People of Color and whites is not the result of a few extreme cases.

ii. In teasing out these statistics, we considered whether, in fact, People of Color were reported by law enforcement to have been suspected of committing more separate criminal offenses in each incident than whites did (which, if true, would raise questions of biased policing and reporting that would warrant further study which lies beyond the scope of this report). We also considered whether prosecutors used additional charges against People of Color more often as a negotiating tool to get them to accept a conviction on a lesser charge and whether whites were not charged for offenses that they could have been. These are questions that cannot be answered without access to the records of the Washtenaw County prosecutor's office.

c. Average Number of Convictions per Case: In 10 of the 11 categories that CREW studied, People of Color also received more convictions per case, on average, than whites.[38] On the low end, People of Color had 6.3% more convictions per case, on average, than whites in suspension of license cases (SLI). The greatest difference in the average number of convictions per case was in homicide cases (HOM) where there was a 77.1% difference between People of Color and whites.

[38] Further analysis must be done to determine whether prosecutors offered different plea deals when charging people for the same felony offenses that more often resulted in whites having their charges dismissed than People of Color.

RECEIVED by MSC 4/12/2022 1:07:49 PM

		Convictions per Case (Avg.)		
		PoC	White	Diff %
FC Cases	ARM	2.3	1.9	21.2%
	AWIM	3.1	2.0	57.3%
	GBH	2.7	1.9	39.8%
	HOM	4.1	2.3	77.1%
FH Cases	ARO	2.4	2.5	-6.0%
	DEL	1.8	1.4	25.8%
	POS	1.8	1.6	16.2%
	SLI	1.7	1.6	6.3%
	WCC	1.9	1.4	31.8%
	WFF	3.5	2.0	74.0%
	WPF	2.7	2.1	25.7%

i. Is the imbalance in the average number of convictions due to the prosecutor or the judge, or both? Most criminal cases do not go to trial; they end with a plea deal whereby the defendant usually agrees to a conviction(s) in return for dismissing other charges or a recommendation for a lighter sentence. But if the prosecutor starts off by filing more charges for People of Color than whites, it is foreseeable that People of Color would end up with more convictions. A defendant facing a higher number of charges could perceive a greater threat and feel compelled to agree to more convictions. While there is typically a blind rotation by which judges are assigned a particular case, prosecutors' decisions related to such things as charges, bail, or pleas continue after the assigned judge is known. Prosecutors might tailor their decisions to the judge before whom they are appearing. For example, if a judge is viewed as giving harsher sentences to a certain group of people or for a certain class of offenses, the prosecutor might include additional charges or might not offer a plea to that group of people or when the charges involve those offenses.

d. Percentage of People charged as a Habitual Offender: In all 11 categories, prosecutors charged People of Color as Habitual Offenders more often than whites. The only category where People of Color and whites were similarly charged as Habitual Offenders was in the suspension of license category where 3% more People of Color were charged than whites. On the other end of spectrum, in Carrying Concealed Weapon cases, prosecutors used the Habitual Offender designation 653% more for People of Color than for whites.

RECEIVED by MSC 4/12/2022 1:07:49 PM

		Cases Charged as Habitual Offender (%)		
		PoC	White	Diff %
FC Cases	ARM	35.0%	25.6%	36.5%
	AWIM	41.5%	10.0%	314.6%
	GBH	31.4%	23.5%	33.6%
	HOM	44.7%	0.0%	-
FH Cases	ARO	27.9%	21.4%	30.2%
	DEL	28.1%	8.7%	223.4%
	POS	32.9%	18.1%	81.8%
	SLI	30.9%	30.0%	3.0%
	WCC	20.4%	2.7%	653.7%
	WFF	44.0%	0.0%	-
	WPF	44.8%	25.0%	79.1%

i. The Habitual Offender (HO) designation is another area of discretion for prosecutors. If a person has previously been convicted of one or more felonies (or attempts to commit felonies), a prosecutor *may* choose – but is not required - to designate that person as a "habitual offender" if charging that person for a subsequent felony charge. If that person is convicted, s/he faces a much longer sentence that could lengthen the prison sentence by 25% - 100%, depending on the number of prior convictions. One might think the HO designation makes sense for prosecutors to use in instances when they are charging someone who displays a serious pattern of ongoing criminal behavior. But that is not the only way it is used.

In 2008, the Michigan Supreme Court said in People v. Gardner that the habitual offender law passed by the legislature was so broad that it required courts to count each separate felony *conviction* in calculating the amount by which a sentence will increase under the habitual offender law, not the number of separate criminal *incidents* resulting in felony convictions. In other words, if someone has more than one conviction stemming from one incident, each conviction would add on additional time to the person's sentence if the prosecutor opts to charge that person as a Habitual Offender in connection with another felony down the road.

That is one of the reasons why the statistics on average number of convictions for People of Color versus whites is so disturbing—more convictions per case makes it more likely a Person of Color will be charged as a Habitual Offender, and receive a significantly longer sentence, if she/he gets charged again with a felony. The likelihood that People of Color will get charged as Habitual Offenders is also magnified when you consider the

RECEIVED by MSC 4/12/2022 1:07:49 PM

rate at which they are charged with a crime and the average number of charges they receive per case.[39]

ii. We were troubled here by two things. First, to see how often Washtenaw County prosecutors charged people as Habitual Offenders overall when the decision whether to charge someone as a Habitual Offender is discretionary. Second, the racial disparity in terms of who was charged as a Habitual Offender. For example, in the homicide cases, almost half of People of Color were charged as Habitual Offender while no whites were charged in that way. Could whites have been charged as a Habitual Offender but were not? Does the breadth of how the HO label is applied indicate use or misuse of the HO tool beyond its intended purpose, and, if so, are discretionary decisions by the Washtenaw County prosecutor's office responsible for a racial disparity?

C. Patterns in Judicial Decision-Making

1. Overview
   a. CREW utilized a sample of 3,608 charges for its initial review of judicial decision-making patterns. The charges analyzed in the four capital felony (FC) case categories represent seven years of judicial dispositions while the charges included in the seven noncapital felony (FH) categories cover three years of judicial dispositions. While we recommend further data review and ongoing transparency, the analysis of all judges' rulings in these felony categories was sufficient to conclude that racial disparity is a concern that must be addressed in the Washtenaw County Circuit Court's criminal proceedings.

      CREW took into consideration that there might be instances where an individual judge's sentencing decision showed disparity in one of the case categories.[40] In fact, our data does offer examples of one judge or another being out of step with others on the court or showing racial disparities in sentencing in a single case category. We highlighted 23 instances across the case categories that we studied where an individual judge was out of step with others on the court in terms of harsher sentencing or whose sentences showed racial disparities.

      Of more concern, though, was whether we would find a pattern of disparities across all 11 case categories by one or more judges or by the entire court. While we found racial disparities among individual judges, CREW did not find that the Washtenaw County Circuit Court, as a whole, demonstrated a pattern of racial disparity in its sentencing of our community members across all case categories we studied.

[39] In addition to possible bias in charging, Blacks being charged more often with Habitual Offender status may reflect other factors, such as heavy law enforcement presence in low-income Black communities.

[40] Judicial discretion is valued and should not be limited artificially; indeed, when it considers appropriate factors, it can enhance justice. It just should not be used a discriminatory manner.

RECEIVED by MSC 4/12/2022 1:07:49 PM

That is not to say, however, we found no instances of disparity across the court. As an example, our data revealed that 44.1% of People of Color convicted of homicide were sentenced to life in prison compared to only 27.3% of White defendants convicted of homicide and most of the judges contributed to a wide racial disparity in the average minimum/maximum sentence in armed robbery cases.

The data did reveal, however, that one Washtenaw County Circuit judge's sentencing decisions appear to reflect a troubling pattern of racial disparities and/or harsher sentencing practices across the case categories and account for 13 of the 23 issues we raise in our discussion below.

b. The data on how judges handled the charges brought by the prosecutors are detailed in two sets of tables (see Tables). The teal colored tables detail the charges for each of the 11 categories of criminal offenses that CREW studied. The orange colored tables, entitled "Sentencing Statistics by Defendant's Race and Sentencing Judge" analyze what happened when someone charged with an offense was convicted of that offense.

For each of the case categories, we looked at the dispositions by an individual judge as well as the totality of the dispositions by the entire bench. In analyzing the statistics on how judges resolved and disposed of the charged offenses, CREW looked at three areas:

i. First, we studied each case category to see whether any individual judge displayed a significant racial disparity in her/his sentencing for the same charged offense. We found instances where there was a significant difference between the length of prison and jail sentences imposed on whites as compared to People of Color. We also found and noted disparities in the length of probation imposed, based on race.[41] In this section of the report, while we highlight certain areas of racial disparity, this is not intended to suggest that the others areas do not require further consideration or analysis.

ii. Second, we assessed whether any of the sitting Washtenaw County Circuit judges were an outlier in terms of the average minimum/maximum prison or jail sentences or probationary term lengths he/she imposed for specific categories of charged cases.

iii. Third, we looked to see whether a racial disparity occurred in terms of the type of punishment imposed by a judge upon conviction among the available options of probation only, a jail sentence or a prison sentence.

c. Of the 11 case categories, there were 3 categories that did not demonstrate significant racial disparities among any of the judges nor was there any outlier in terms of sentencing ranges.

[41] CREW did not have the capacity to review the terms of probation, other than the length, to assess any further racial disparities in punishment regarding probation terms including restrictions and/or costs imposed. We did collect that data which could be studied in a subsequent review.

RECEIVED by MSC 4/12/2022 1:07:49 PM

These areas have narrow statutory sentencings that limit judicial discretion in sentencing. The areas were:

i. Suspended License (SLI) – a category for which there is a high dismissal and low conviction rate.
ii. Weapons Felony Firearm (WFF) – a category for which the judge has no discretion in terms of the mandatory sentence. Only 2 WFF charges were filed by Washtenaw County prosecutors against whites and both of those charges were dismissed. At the same time, 64% of People of Color were convicted of WFF. The prosecutors also designated 44% of People of Color charged with felony firearm as a Habitual Offender while none of the white people charged were designated as a Habitual Offender.
iii. Homicide (HOM) – a category for which sentencing discretion is limited. For example, there is no discretion in a sentence for a first-degree homicide conviction for an adult. The law requires a mandatory life without parole sentence.

d. Of the remaining 8 case categories, Washtenaw County Judge Brown's sentencing decisions raised at least one issue in all categories. Those issues were significant racial disparities in sentencing, harsher average sentences that were outliers compared to the rest of the court or significant racial disparities with regard to where whites versus People of Color carried out their sentences (e.g. probation versus prison). These disparities were greater and more consistent as compared to data regarding the cases before other judges.

The summary table below provides an overview of the case categories discussed in this section and notes where we found instances of racial disparity or harsher sentencing practices among the judges. Each instance is discussed further throughout this section.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Offense	Judge Brown	Judge Kuhnke	Judge O'Brien	Judge Swartz
ARO	*Disparity/Place			*Disparity/Sen *Disparity/Place
DEL	*Disparity/Sen *Disparity/Prob *Disparity/Place			*Disparity/Place
POS	*Disparity/Sen			
WCC	*Disparity/Sen *Disparity/Place	*Disparity/Prob		*Disparity/Place
WPF	*Disparity/Sen *Disparity/Prob *Disparity/Place	*Outlier		
ARM	*Disparity/Sen		*Disparity/Sen	*Disparity/Sen *Outlier
GBH	*Disparity/Sen			*Disparity/Sen
AWIM	*Outlier			

Key: Racial Disparity in Sentencing = Disparity/Sen
Racial Disparity in Placement = Disparity/Place
Racial Disparity in Length of Probation = Disparity/Prob
Outlier/Higher Minimum and Maximum Sentence = Outlier

2. Non-Capital Felony Cases (FH Tables)

a. Assaulting, Resisting, or Obstructing an Officer (ARO): This is one of the most frequent charges made in Washtenaw County among the FH cases and is almost always an add on charge to another felony.

i. Disparity in Sentencing

- The percentage of whites convicted of ARO was far lower than People of Color. Though there were a limited number of whites convicted, in general, whites received a slightly higher number of years when sentenced to prison among three of the judges.
- There was a wide difference in the average minimum/maximum prison sentence that Judge Swartz gave to People of Color (1.5 – 4.1 years) versus whites (.4 – 2 years).

RECEIVED by MSC 4/12/2022 1:07:49 PM

ARO Sentence Disparity	Judge Swartz		
	PoC	White	% Diff.
Average Min. Prison Sentence (yrs.)	1.5	0.4	250.0%
Average Max. Prison Sentence (yrs.)	4.1	2.0	106.3%

ii. Differences in Where Defendants Served their Sentence

- For two of the judges there was a disparity based on race as to who received the lesser punishment of probation for a conviction of this charge. Judges Brown and Swartz sentenced the largest share of their white defendants to probation (42.9% and 50%, respectively) while sentencing the largest share of People of Color to incarceration in jail (52% and 56.3%, respectively).

ARO Placement	Judge Brown			Judge Swartz		
	PoC	White	% Diff.	PoC	White	% Diff.
Sentenced to Jail (%)	52.0%	28.6%	82.0%	56.3%	25.0%	125.0%
Sentenced to Probation Only (%)	16.0%	42.9%	-62.7%	18.8%	50.0%	-62.5%

b. Controlled Substance-Delivery (DEL): Most convictions for this charge result in jail or probationary sentences.

i. Disparity in Sentencing

- Where jail or prison sentences were imposed, Judge Brown gave People of Color significantly longer jail sentences than whites (0.6 years versus 0.2 years.

DEL Sentence Disparity and Probation Length	Judge Brown		
	PoC	White	% Diff.
Average Jail Sentence (yrs.)	0.6	0.2	290.0%
Average Probation Period Accompanying Jail Sentences (yrs.)	2.17	1.00	116.7%

- Judge Brown gave People of Color longer terms of probation after release from jail (2.17 years) than he gave to whites (1 year). [42]

ii. Differences in Where Defendants Served their Sentence

- No one on the bench sentenced a white person convicted of this charge to a prison sentence. As a group, the court gave the least harsh punishment –

[42]Longer terms of supervision correlate with an increased likelihood of a violation of one of the terms of probation often resulting in incarceration.

RECEIVED by MSC 4/12/2022 1:07:49 PM

probation only – 66.7% of the time to white defendants. Only 31.8% of People of Color received probation while the rest were incarcerated either in jail or prison.

- Judge Swartz placed 100% of the whites convicted of DEL on probation while placing 0% of the People of Color on probation. People of Color all received prison or jail sentences.
- Judge Brown placed 2/3 of white defendants on probation while sentencing the majority of People of Color to incarceration in prison or jail.

DEL Placement	All Judges			Judge Brown			Judge Kuhnke			Judge O'Brien			Judge Swartz		
	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.
Probation Only Cases (%)	31.8%	66.7%	-52.3%	41.2%	66.7%	-38.2%	66.7%	66.7%	0.0%	10.0%	0.0%	-	0.0%	100%	-

c. Controlled Substance – Possession (POS)

i. Disparity in Sentencing

- Judge Brown's average minimum/maximum prison sentences were significantly higher for People of Color (1.5 – 8.5 years) than whites (1 – 3.5 years).

POS Sentence Disparity	Judge Brown		
	PoC	White	% Diff.
Average Min. Prison Sentence (yrs.)	1.5	1.0	53.2%
Average Max. Prison Sentence (yrs.)	8.5	3.5	142.9%

d. Weapons- Carrying Concealed (WCC)

i. Disparity in Sentencing: Judge Brown's average minimum/maximum prison sentences were significantly higher for People of Color (2.4 – 6 years) versus whites (.5 – 5 years).

WCC Sentence Disparity	Judge Brown		
	PoC	White	% Diff.
Average Min. Prison Sentence (yrs.)	2.4	0.5	372.2%
Average Max. Prison Sentence (yrs.)	6.5	5.0	30.0%

ii. Differences in Where Defendants Served their Sentence

- The entire bench displayed a large difference in the percentage of people convicted whom they sent to prison. They incarcerated a significantly higher percentage of People of Color in prison than whites.

RECEIVED by MSC 4/12/2022 1:07:49 PM

- Judge Brown gave probation in 90% of the cases involving whites and only 54.5% of the cases involving People of Color.
- Judge Swartz gave probation in 88.9% of the cases involving whites and only 44.8% of the cases involving People of Color.
- Judge Kuhnke gave People of Color and whites just probation at roughly the same rate but the length of probation was significantly higher for People of Color than whites (3 years versus 1.5 years).

WCC Placement	All Judges			Judge Brown			Judge Kuhnke			Judge O'Brien			Judge Swartz		
	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.
Prison Cases (%)	14.6%	2.9%	410.4%	13.6%	10.0%	36.4%	13.0%	0.0%	-	13.6%	0.0%	-	17.2%	0.0%	-
Probation Only Cases (%)	39.6%	62.9%	-37.0%	54.5%	90.0%	-39.4%	56.5%	62.5%	-9.6%	0.0%	0.0%	-	44.8%	88.9%	-49.6%

WCC Disparity Probation Length	Judge Kuhnke		
	PoC	White	% Diff.
Average Probation Only Period (yrs.)	3.0	1.5	102.6%

e. Weapons – Felon in Possession (WPF)

i. Disparity in Sentencing: There were significant racial disparities in Judge Brown's average minimum/maximum prison sentence for People of Color (2.7 - 7.2 years) versus whites (.6 – 5 years).

ii. Outliers in Min/Max Sentencing: Judge Kuhnke's average minimum/maximum sentences were significantly higher than the rest of the court.

WPF Sentence Disparity & Outlier	Judge Brown			Judge Kuhnke		
	PoC	White	% Diff.	PoC	White	% Diff.
Average Min. Prison Sentence (yrs.)	2.7	0.6	361.9%	5.3	-	-
Average Max. Prison Sentence (yrs.)	7.2	5.0	43.3%	16.1	-	-

iii. Differences in Where Defendants Served Their Sentence

- Judge Brown gave the lesser punishment of probation to 75% of the whites convicted of WPF while giving probation to 22.2% of People of Color.
- When he did give People of Color the lesser punishment of probation, Judge Brown handed out much longer probation terms for them than he did for whites (2.5 years versus 1.7 years).

RECEIVED by MSC 4/12/2022 1:07:49 PM

WPF Probation Length & Placement	Judge Brown		
	PoC	White	% Diff.
Probation Only Cases (%)	22.2%	75.0%	-70.4%
Average Probation Only Period (yrs.)	2.5	1.7	50.0%

3. Capital Felony Cases (FC Tables)

a. Armed Robbery

i. Disparity in Sentencing: Our data shows a racial disparity in sentencing of cases involving an armed robbery conviction for the three judges who sentenced People of Color and whites for armed robbery: Brown (132.6% difference in average minimum sentence between whites and People of Color), O'Brien (115.1% difference) and Swartz (73.5% difference).

ii. Outliers in Min/Max Sentencing: Judge Swartz's average min/max sentence for People of Color was noticeably higher than the other judges making it an outlier on the court. The court as a whole had an average of 9.8 year min/23.7 year max for People of Color while Judge Swartz averaged 14.1 year min/31.0 years max.

ARM Sentence Disparity & Outlier in Sentence Length	All Judges			Judge Brown			Judge Kuhnke			Judge O'Brien			Judge Swartz		
	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.
Average Min. Prison Sentence (yrs.)	9.8	4.9	101.9%	9.7	4.2	132.6%	6.1	-	-	10.1	4.7	115.1%	14.1	8.1	73.5%
Average Max. Prison Sentence (yrs.)	23.7	18.9	25.6%	24.9	19.9	24.8%	17.7	-	-	23.6	16.0	47.4%	31.0	22.0	41.1%

b. Assault with Intent to do Great Bodily Harm Less Than Murder (GBH)

i. Disparity in Sentencing

- Judge Brown sentenced whites convicted of GBH to markedly lower average prison sentences than he did People of Color. His average minimum/maximum prison sentence for People of Color was 7.1 - 18.7 years while the average min/max sentence for white people was 2.5 - 10 years.

RECEIVED by MSC 4/12/2022 1:07:49 PM

- Judge Swartz sentenced whites convicted of GBH to markedly lower average prison sentences than he did for People of Color. Judge Swartz's average minimum/maximum prison sentence for People of Color was 8.1 – 21 years while the average minimum/maximum sentence for whites was 3.5 – 12.5 years.

GBH Sentence Disparity	Judge Brown			Judge Swartz		
	PoC	White	% Diff.	PoC	White	% Diff.
Average Prison Min. Sentence (yrs.)	7.1	2.5	185.6%	8.1	3.5	134.3%
Average Prison Max. Sentence (yrs.)	18.7	10.0	86.7%	21.0	12.5	68.0%

c. Assault with Intent to Commit Murder (AWIM)

i. Disparity in Sentencing: A conviction for AWIM carries one of the most severe punishments in criminal sentencing falling just behind a homicide conviction. Unlike a homicide charge where the death of a person at the hands of another person points a prosecutor directly to a homicide charge, an injury of a person at the hands of another person does not necessarily point a prosecutor to a charge of assault with intent to commit murder. There are other assault charges ranging in severity of punishment with which a prosecutor could choose to charge a person who injured another person (e.g., misdemeanors such as assault and battery and felonies such as assault with a dangerous weapon and assault with intent to do great bodily harm less than murder.) An AWIM charge hinges on a subjective element, whether the person charged *intended to murder* another person. Therefore, when CREW saw the wide gap between the number of People of Color and the number of whites charged with AWIM (41 and 10, respectively), it raised questions as to the reasons for choosing the most severe charge possible for injuring another person and whether the outcomes for those charged with AWIM differed depending on one's race.

The data for the period we studied shows that the outcomes for those charged with AWIM do indeed differ in Washtenaw County depending on one's race. During the period CREW studied (2013-2019), 10 whites were charged with AWIM and only 2 were convicted of AWIM. What happened to the others? Judges approved plea deals for those white defendants to the much less severe charges of assault by strangulation, assault with a dangerous weapon and assault with intent to do great bodily harm less than murder. Most of the white defendants received penalties such as probation up to about 2 years in prison.

In contrast, the teal AWIM table in the Table Section shows that People of Color were convicted of AWIM (not a lesser charge) at a much higher rate. In Washtenaw County, if a

RECEIVED by MSC 4/12/2022 1:07:49 PM

Person of Color was charged with AWIM during the time frame CREW studied, that person had a far greater chance of being convicted for AWIM than if a White person was charged.

The disparity is most striking in data pertaining to convictions in cases presided over by Judge Brown. Almost 2/3 of People of Color charged with AWIM in cases presided over by Judge Brown were convicted of AWIM and none of the whites who came before him charged with AWIM were convicted of AWIM. To put it another way, 36.4% of People of Color charged with AWIM did not get convicted of AWIM before Judge Brown while 100% of whites charged with AWIM did not get convicted of AWIM before him. None of the other judges convicted People of Color of AWIM at such a high rate (63.6%) as Judge Brown. Judges Kuhnke, O'Brien and Swartz convicted 14.3%, 21.4% and 11.1% of People of Color of AWIM, respectively.

AWIM Sentence Disparity	All Judges			Judge Brown			Judge Kuhnke			Judge O'Brien			Judge Swartz		
	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.
Cases Charged with AWIM	41	10	-	11	4	-	7	1	-	14	2	-	9	3	-
Cases Convicted of AWIM	12	2	-	7	0	-	1	0	-	3	1	-	1	1	-
Cases Convicted of AWIM (%)	29.3%	20.0%	46.3%	63.6%	0.0%	-	14.3%	0.0%	-	21.4%	50.0%	-57.1%	11.1%	33.3%	-66.7%

ii. Outliers in Min/Max Sentencing: Judge Brown also gave People of Color convicted of AWIM significantly longer average minimum sentences (almost 21 years) compared to Judge O'Brien (11 years) and Judge Swartz (7 years).

AWIM Outlier	All Judges			Judge Brown			Judge Kuhnke			Judge O'Brien			Judge Swartz		
	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.	PoC	White	% Diff.
Average Prison Min. Sentence (yrs.)	17.0	10.0	70.0%	20.7	-	-	-	-	-	11.0	10.0	10.0%	7.0	-	-

To understand the picture that emerged from studying Judge Brown's sentencing disparities in all case categories discussed throughout this section of the report, CREW decided to review Judge Brown's AWIM cases going to 2005. We wanted to take extra care and

RECEIVED by MSC 4/12/2022 1:07:49 PM

provide more data believing that a larger dataset would clarify whether the disparities we were seeing were anomalies or part of a longstanding pattern.

The table below shows the pattern of disparity in Judge Brown's sentencing of AWIM charges since 2005. There were 36 cases of people charged with AWIM and convicted of one or more charges in front of Judge Brown during that period. As in the 2013-2019 review, no white person before Judge Brown was convicted of AWIM. The average minimum sentence Brown gave to People of Color convicted of AWIM is similar to his average minimum sentence in the original time frame of 2013-2019 that CREW studied (20.7 vs 20.1 years).

We can see the outcomes for those charged with AWIM differ depending on one's race. Overall, People of Color (as noted earlier over 98% of whom were Black people) charged with AWIM received, on average, a minimum sentence of 11 years compared to a minimum sentence of 4 years for White people charged with AWIM.

Convictions for Cases with AWIM Charges Filed Between 2005-2019, Excluding Life Sentences: Judge Archie Brown	Charged				Charged & Convicted			
	PoC	White	Difference		PoC	White	Difference	
Assault With Intent to Murder								
Number of Cases	24	12	Number	Percent	11	0	Number	Percent
Average Minimum Sentence (Years)	11.0	4.0	6.9	172%	20.1	-	-	-
Average Number of Prior Convictions	1.5	1.0	0.5	50%	1.5	-	-	-
Average Number of Charges (Current Case)	7.0	5.9	1.1	18%	8.3	-	-	-
Average Number of Convictions (Current Case)	3.6	2.8	0.8	30%	5.3	-	-	-

V. CREW'S INTERPRETATIONS AND RECOMMENDATIONS

CREW's work operates under the belief that the institutions that form the basis of our criminal legal system must not only operate in a non-discriminatory manner in fact but must also be perceived to operate in that way. That belief drives the recommendations set forth below.

The data CREW compiled sheds light on deep racial disparities within Washtenaw County's criminal legal system that members of the public have undoubtedly experienced but never seen in documented form. In this Section, CREW puts forth recommendations for action and mechanisms to address racial inequities and discriminatory treatment in our criminal legal system. Discriminatory treatment in the criminal legal system permanently alters a community member's ability to participate equally in our

RECEIVED by MSC 4/12/2022 1:07:49 PM

society.[43] Section VI (Additional Analysis) lays out other areas of review and analysis with the goals of better defining and understanding the extent and cause of the disparities and, importantly, rectifying them.

A. Recommendations: Prosecution

The door that the statistics opens for us shows a disturbing pattern in prosecutions in Washtenaw County. The data illustrates the breadth and depth of a potentially problematic approach to justice in our county, in that disturbing racial disparities are evident across the board. At the very least, the data raises questions (although it does not answer them) about the exercise of prosecutorial discretion, whether there are cultural norms or implicit biases within or outside of the prosecutor's office that are contributing to unfairness, and whether and what kind of systemic reforms are necessary. To address the deeper issues, we recommend:

1. Rigorous financial and administrative oversight by the Washtenaw County Board of Commissioners, including the establishment of a citizen's race equity commission chosen with input of community members.
2. The Board of Commissioners, in collaboration with the prosecutor-elect and the race equity commission, engage a neutral, third-party evaluator to study prosecutors' files, policies and procedures and make recommendations for data driven, evidence-based improvements to rectify racial disparities and determine whether the tools and practices employed by the Washtenaw County prosecutor's office are applied in way that is not racially discriminatory and does reflects the fair administration of justice. Issues to review and/or revise include, among other things:
    a. The disparate impact that the Habitual Offender designation has on People of Color.
    b. The charging of Weapons Felony Firearm. Are there written guidelines; if so, are they applied without regard to race?
    c. The factors used to offer plea deals and the creation of a transparent process for plea negotiations.
    d. The impact that the race of the victim might have on such things as charging decisions, conviction rates, and sentencing decisions.
    e. The use of vertical and horizontal charging practices.
    f. The factors/circumstances involved in deciding to dismiss charges including the cases where all charges are dismissed.
    g. Current use of diversion (i.e., before and after filing charges) and restorative justice; the implementation of evidence-based diversion, deflection and/or restorative justice practices that are shown to improve outcomes for people involved in the County's criminal legal system in a racially neutral manner.
    h. The current performance review process for assistant prosecutors and other staff to ensure it is consistent with the racial equity considerations outlined in this report and supports the "minister of justice" role that prosecutors should play in the legal system.

[43] Beyond the obvious impact of loss of freedom by incarceration, felony records impose disabilities and barriers on access to housing, employment, and education which drastically limit returning citizens' ability to successfully participate in our community.

RECEIVED by MSC 4/12/2022 1:07:49 PM

i. All assault arrests to identify how decisions are made regarding which charges to file (i.e., misdemeanor, felony or no charges) and whether People of Color are treated the same as whites in terms of the severity of the assault charges filed.
j. Community members' experiences in the criminal legal system.

3. The creation of an online dashboard to display data points like those documented by CREW so the public can assess whether the decisions by the Washtenaw prosecutor's office has a disparate impact on racial minorities. Citizens elect the prosecutor but have almost no data upon which to assess the performance of the office; a publicly available dashboard would provide voters and others with such data and improve the transparency of justice. A dashboard would also assist the prosecutor's office because they do not maintain aggregated data on data points such as race, age, and income.[44]

4. Developing written policies and procedures for charging decisions made within the Washtenaw County prosecutor's office. Any such policy/procedure should be accompanied by a racial impact statement (i.e. an analysis assessing the possible racial consequences of any proposed policy before adopting it to avoid any unintended disparate racial effects.)[45]

5. Developing and implementing transparent mechanisms, processes and/or rules for exercising discretion, including a robust process to review assistant prosecutors' decisions regarding cases to ensure they are not engaging in racially weighted decisions that have a disparate impact on racial minorities.

6. Implementing a mandatory, ongoing training program on implicit and explicit racial bias for all county employees within the prosecutor's office and those outside the office used to investigate cases.

7. Developing and implementing regular equity audits to ensure that county employees within the prosecutor's office and individuals and entities used by that office to investigate cases are not engaging in practices that have a disparate impact on racial minorities in Washtenaw County.

8. The Michigan Legislature use CREW's data on the discriminatory impact of the Habitual Offender law as well as available research to eliminate the Habitual Offender statute.

B. Recommendations: Judicial Decision-Making

1. Since state courts across the country have acknowledged that the judicial system can be part of the problem in the unequal treatment of minorities, particularly with regard to Black people in the criminal legal system, and that such discriminatory treatment contributes to mass incarceration of People of Color, we are compelled to ask why hasn't our court developed any method of transparency in reporting these problems to the public and initiated oversight and change? Since

[44] In answer to a FOIA request asking whether the Washtenaw County prosecutor's office has aggregated data on such things as age and race, First Assistant Prosecutor Eric Gutenberg wrote, "This office does not have aggregated data on data points such as race or age of those charged." Email dated August 17, 2020 from Eric Gutenberg to MaryAnn Sarosi.

[45] In fact, any policy or procedure adopted regarding prosecutorial functions should be accompanied by a racial impact statement as states such as Iowa, Connecticut, Oregon, and New Jersey have done. *Report to the United Nations on Racial Disparities in the U.S. Criminal Justice System*, The Sentencing Project, 2018 https://www.sentencingproject.org/publications/un-report-on-racial-disparities/

RECEIVED by MSC 4/12/2022 1:07:49 PM

our case data came from the court's own public website, but in a format that takes hundreds of hours to collect and analyze, it's natural for us to wonder why the court hasn't performed an equity audit in the past; or, if it has, why such an audit hasn't been made public so voters can make an informed decision when they elect a judge. While the disparities revealed in our analysis pre-date the tenure of the current chief judge of the Washtenaw County Court, we strongly encourage the use of our report to assist the current court to undertake serious and transparent reforms.

2. The Washtenaw County Board of Commissioners also has an integral role to play in the reforms needed to address racially discriminatory practices in the county's criminal legal system. Not only does the Board have an interest in ensuring that the county's criminal legal system aligns with the *One Community: Advancing Racial Equity in Washtenaw County* initiative, it has oversight of the county's budget, 20% of which is allocated to judicial functions.[46] Given the responsibility of the Board as well as the Washtenaw County Court to ensure equal justice in the criminal legal system, they could jointly establish a citizens race equity commission made up of community members appointed by relevant stakeholders to:
    a. Annually compile, study, and publish statistics such as those analyzed by CREW to detect racial disparities in the Washtenaw County Courts. The data could be posted on an online dashboard described below.
    b. Address the issues and questions raised in Section VI of CREW's report. As noted in that section, some of the questions can be answered using the capital felony (FC) and non-capital felony (FH) dataset that CREW has already compiled while other questions require additional information that was not available to CREW. For example, the racial equity commission could examine the cases involving life sentences to understand why life sentences are given to People of Color at such a disproportionate rate compared to whites.

3. CREW noted 23 instances where a Washtenaw County judge's sentencing decision exhibited racial disparity or was an outlier in terms of handing down harsher sentences than others on the court. More than half of the concerns CREW raised in this report (i.e., 13 of the 23) come from the judicial decision-making of one judge. The other 3 judges combined comprised less than half of our concerns. With those 23 instances in mind, we urge the Washtenaw County Circuit Court to, among other things:
    a. Study the 23 issues further by gathering additional information on the cases involved to understand and address the source of the disparities. The study could be done under the auspices of the citizens race equity commission described above and would address the questions raised in Section VI (Additional Analysis) below.
    b. Study whether there is an imbalance among the judges in terms of defendants designated as Habitual Offenders, average number of convictions, percentage of dismissed charges, etc.

4. The disparities in cases presided over by Judge Brown are so strikingly consistent among all eight categories we highlighted and affect so many people that the court should consider taking

[46] Washtenaw County 2019-2022 Preliminary Budget Summary, D-4 https://www.washtenaw.org/DocumentCenter/View/11250/2019-2022-Preliminary-Budget-Summary

RECEIVED by MSC 4/12/2022 1:07:49 PM

immediate action (in collaboration with an impartial citizens race equity commission). Therefore, we:

a. Suggest that the Washtenaw County Circuit Court engage a neutral, outside, third-party to determine whether the racial disparities evidenced in Judge Brown's sentencing decisions are rooted in any personal bias or systemic bias. To fully capture and understand the source of the disparities, the third-party entity should be permitted and encouraged to: thoroughly review the cases filed in the case category where the issue is raised in our report during the time frame CREW studied; notify the defendants in the cases under review of the existence of CREW's report as well as the review underway by the third-party entity; solicit confidential interviews with party participants in the cases and solicit public comment from those who have been involved and impacted; and inquire whether additional information is needed to assess the disparities in question.
b. Believe the review should be completed in a timely manner in the interest of justice.
c. Understand that the public could rightfully ask for assurance that the pattern of disparities in Judge Brown's sentencing decisions do not continue during the review period and believe the court has a variety of tools it could implement to instill confidence in the justice system. For example, the Washtenaw County Circuit Court could supervise/monitor his criminal cases, place a moratorium on sentencing of Judge Brown's pending criminal cases or shift his criminal caseload temporarily pending outcome of the review. Taking any of those steps would signal to the public that the court understands that judges must be neutral in fact and be publicly perceived "as an impartial dispenser of justice."[47]

5. The Washtenaw County Court could also:
   a. Issue a Local Administrative Order addressing racial bias and prejudice in conduct and judicial decisions such as sentencing.
   b. Institute a mechanism for receipt of public comment and complaints related to bias and prejudice in judicial conduct (including judicial staff) and judgments with a mechanism for investigation, review, and implementation of any remedies.
   c. Develop and implement a publicly available dashboard that is searchable and updated regularly for reporting such things as: all charges, all sentences, pleas with details of offense, and race broken down by judge. The creation of an online dashboard by the Washtenaw County Court and/or the State Court Administrative Office will assist the public in assessing whether a judge's decisions have a disparate impact on racial minorities.
   d. Establish a protocol whereby all policy and procedures adopted by the Washtenaw County court regarding judicial functions in the criminal legal system should be accompanied by a racial impact assessment.
   e. Provide ongoing equity/implicit bias training to all county employees at the courthouse and ensure that their performance reviews reflect any concerns about unequal treatment of courthouse users. We recognize that judges are not the only tax-funded court employees whose decision could result in disparate treatment and outcomes. Judicial attorneys, for example, participate in decision-making and draft rulings in cases and can, therefore,

[47] *People v. Killebrew*, 416 Mich. 189, 202; 330 N.W.2d 834 (1982). Such a review would not preclude a person from filing a complaint with the Judicial Tenure Commission or taking similar action.

RECEIVED by MSC 4/12/2022 1:07:49 PM

impact proceedings. As with judges, they should not administer justice in a preferential manner whether in fact or by perception.[48] When a court employee's actions suggest unequal treatment of a person/persons using the court, action should be taken so Washtenaw County residents can have confidence that everyone receives equal justice under law.

6. Oversight of Racial Disparities in Judicial Decision-Making

Judges are human and not necessarily immune from the myths and biases, implicit as well as explicit, that have informed decisions and impacted our criminal legal system laws for decades, resulting in racial disparities in arrests, treatment, and incarceration of People of Color at astonishingly high rates.[49] As CREW studied the data on patterns of racial disparities related to sentencing decisions by Washtenaw County judges, we looked to see whether there was an oversight mechanisms that should have caught the disparities and possible discriminatory treatment evident in our data.

First, we looked at the court of appeals. While there is appellate review of judicial sentencing and decisions, CREW's understanding is that the court of appeals limits its review to whether there is abuse of discretion or error in applying the law to the facts of an individual case. There appears to be no effective mechanism for appellate review of overall patterns of racial disparities that could indicate discriminatory or racially biased rulings.

Second, we looked to the ethical canons embodied in the Code of Judicial Conduct. While the Michigan Supreme Court authorizes the Canons, adherence to the Canons is supervised by the Judicial Tenure Commission (JTC) which is responsible for "investigating complaints of judicial misconduct and judicial incapacity, and for recommending discipline of judges by the Michigan Supreme Court."[50]

The JTC, however, is limited in its review and investigation to conduct that violates the Code of Judicial Conduct, and Michigan's Code does not explicitly speak to racial bias or prejudice in decision-making, including that which may result in racial disparities. The American Bar Association's Model Code of Judicial Conduct, by contrast, includes a Canon explicitly prohibiting racial and gender prejudice in performance of a judge's duties. Other states have adopted such a Canon for judges, such as Canon 3 B (5) of the Virginia Canons of Judicial Conduct, which states that:

---

[48] A judicial attorney to one of the Washtenaw County judges suggested in a news article that she may not treat people equally when she said, "I really do keep a list of attorneys that annoy me. (There are currently 9 people on the list and you know who YOU are.)" While her comment might have been intended as a joke, it might not be perceived that way by the public and serves to undermine confidence in the justice system. http://www.legalnews.com/washtenaw/1266844

[49] Sara Beale, *The News Media's Influence on Criminal Justice Policy: How Market-Driven News Promotes Punitiveness,* 48 Wm. & Mary L.Rev. 397(2009); John J. Dilulio, Jr., *My Black Crime Problem and Ours: Why Are So Many Blacks in Prison? Is the Criminal Justice System Racist? The Answer is Disquieting,* City Journal, Spring 1996*; Hon. Harold Hood, The Race/Ethnic Bias Task Force Four Years Later—Looking Back,* 73 Mich B.J.267 (1994)

[50] The authority and composition of the membership of the Judicial Tenure Commission comes from Article 6 Sec 30 of the Michigan Constitution. Does the composition of the JTC which has been in place since 1968 (where 5 of the 9 members must be judges, 2 more are lawyers and only the remaining 2 are non-lawyers) adequately reflect input from the public or serve as an effective oversight body such that the public can have confidence in its fairness?

RECEIVED by MSC 4/12/2022 1:07:49 PM

"A judge shall perform judicial duties without bias or prejudice. A judge shall not, in the performance of judicial duties, by words or conduct manifest bias or prejudice, including but not limited to bias or prejudice based upon race, sex, religion, national origin, disability, age, sexual orientation or socioeconomic status, and shall not permit staff, court officials and others subject to the judge's direction and control to do so.[51]

Given that Michigan's Canons do not explicitly cover bias and prejudice[52], we are concerned that racial disparities in sentencing People of Color may not receive the attention it deserves by the JTC. In fact, the JTC expressly states that claims involving a judge's discretionary handling of judicial duties do not fall under misconduct and is, therefore, not within the JTC's authority.[53] That leaves open the question, "Where does oversight lie for a judge's discretionary handling of judicial duties, when such acts include racial disparities in sentencing?"

Third, we considered if the public has oversight capabilities to assess whether a judge's decision-making results in racial disparities to the disadvantage of People of Color. In Washtenaw County, and indeed across Michigan, the public lacks effective mechanisms for review of a judge's record as there are no audits, scorecards or dashboards to inform the public of disparities and other patterns that emerge from the decisions of a specific judge.

Without such a mechanism, there is no basis for the public to make fully informed election decisions regarding incumbent judges, when racial disparities in decision-making or signs of bias or prejudice are important to voters. So, while the voters in Michigan elect judges, the public is largely flying blind.

Thus, CREW was left with the same question with which we started: How should judges in Michigan be held accountable for patterns of discriminatory judgements and/or biased treatment of individuals who come to the court?

To answer this question, we urge the Michigan Supreme Court to:

a. Examine the ways in which the Michigan Constitution protects against racial bias, discrimination, and prejudice and how those protections are evidenced in the criminal legal system.

b. Issue for public comment, a Canon that is consistent with the ABA Model Judicial Canons, which explicitly addresses the mandate that a judge's conduct (and that of

[51] http://www.courts.state.va.us/courts/scv/canons_of_judicial_conduct.pdf

[52] Michigan Canon 3(A)(14) does provide: "Without regard to a person's race, gender, or other protected personal characteristic, a judge should treat every person fairly, with courtesy and respect." However, the Canon does not explicitly address bias or prejudice in decision-making, including that which may result in racial disparities.

[53] State of Michigan Judicial Tenure Commission Annual Report 2019 at Page 7
http://cms4.revize.com/revize/mjtc/annual_report/docs/2019%20Annual%20Report.pdf

RECEIVED by MSC 4/12/2022 1:07:49 PM

her/his judicial staff) and judgments be without bias and prejudice based on all protected categories and status.

c. Provide a basis for review and audit of racial disparities in criminal sentencing and authorize the JTC to investigate such disparities for violation of the new Canon.

7. We urge the Michigan Judicial Tenure Commission to provide a mechanism for public complaints on racial disparity and bias in judicial conduct and judgements and make such data public.

VI. ADDITIONAL ANALYSIS

A. Additional Analysis Using CREW's Dataset

1. In calculating the sentencing tables (orange tables), we excluded sentences where the defendant received a life sentence or was designated as a Habitual Offender 4th offense. For those given a life sentence, we could not calculate what term of years to use when calculating the average maximum sentence, so we excluded them from calculating sentence lengths reported in the tables. A defendant convicted as a HO 4th offense, would receive a 100% longer sentence which might skew the average maximum sentence for a judge. Although we did not incorporate those 2 factors into the sentencing calculations, there remain significant questions that need to be addressed, namely:

    a. Of the 265 capital felony cases that CREW analyzed, 50 cases had homicide charges (45 of which resulted in a homicide conviction). Of the 45 convictions, 27 were sentenced to non-life sentences and 18 were given life sentences. People of Color made up 83.3% of those 18 people sentenced to life in prison. In other words, 44.1% of People of Color convicted of homicide were sentenced to life in prison, compared to only 27.3% of White defendants convicted of homicide. Further study needs to be done to get to the root of why People of Color (primarily Black people) are sentenced to life in prison in such a disproportionate manner in Washtenaw County.
    b. People of Color are charged as Habitual Offenders far more often than whites; this has a multiplier effect on the length of the final sentence. Is this a form of bias or does this reflect an objective difference in prior convictions?
    c. To what extent does the Washtenaw County prosecutor's office apply the HO designation when a defendant's prior felony convictions result from one event rather than serial criminal behavior?
    d. To what extent is the prosecutor's power to designate a defendant as a Habitual Offender used (either at the outset or during a case) as a bargaining chip for a defendant to accept a plea deal involving a conviction? Is a HO designation dropped more often for whites?

RECEIVED by MSC 4/12/2022 1:07:49 PM

2. Of the 12.6% of the capital felony cases in our dataset in which all charges were dropped and a defendant was not convicted of any charge, People of Color made up 76.8% of this group, whites made up 12.5% and 10.7% of the cases had no race information for the defendants. In the FH study, there were 7.7% of all FH cases in which the defendant was not convicted of any charge. Of that group, People of Color accounted for 55.7% of this group, whites accounted for 41.4% and the race of 2.9% of the defendants was unknown. This is not an insignificant number of cases.

    a. What law enforcement or prosecutorial practices account for the disproportionate number of People of Color (mainly Black people) who were arrested, charged, and then released at some point with no charges being brought? What explains charging these people only to release them? How much taxpayer dollars are being spent arresting, charging, and holding these people? Even if all charges are dismissed and they return home, arresting and charging people impacts their lives in a variety of ways. For example, they are usually incarcerated in jail for some period before the charges are dropped and, in that period of incarceration, they may have lost their job, housing and custody of their children. Additionally, they have been fingerprinted, required to give a DNA sample, and photographed. Their information is then placed in various law enforcement databases.
    b. Further examination of all capital felony (FC) and non-capital felony (FH) cases (beyond the case categories and time frame we examined) could be done to study the percentage of charges that are dismissed and to calculate the costs involved in bringing charges that are routinely dismissed. In other words, what are the costs to Washtenaw taxpayers and individuals who are charged by prosecutors when those charges are ultimately dismissed (i.e. the cost to the person charged as well as the cost of prosecutors, law enforcement and the courts).

3. What accounts for the differences in the average number of charges brought by prosecutors against People of Color versus whites? Are the number of initial charges compared to the number of final charges different for People of Color and whites? If so, what accounts for the difference; could it indicate that prosecutors are filing some charges with the intention of using them to drive defendants toward agreeing to a conviction on lesser charges and not intending to prove the elements of the higher charges (a practice known as vertical overcharging)?

4. Further study should be done to assess whether there are differences in the facts of cases such prior criminal history, age, weapon, race of victim or location of the incident that could account for the appearance of such a severe disparity in punishment based on race.

5. Would examining fines levied indicate possible racial disparities?

RECEIVED by MSC 4/12/2022 1:07:49 PM

6. Would comparing charges, convictions and sentencing by age show racial disparities, e.g., are young People of Color and young White defendants treated similarly for similar behavior? National data indicates that there are racial disparities among young people. [54]

7. Is there data that would show why far fewer white people are charged in some areas (e.g. AWIM, weapons cases)? Are they not being arrested, not being charged, or charged for lower offenses?

8. Would comparing the length of time between case milestones, e.g., filing date, disposition date and sentencing date, show disparities in terms of length of time between milestones? For example, are People of Color subjected to longer periods in jail before they are offered a plea.

9. Would data on the type of representation show disparate outcomes for a case with a retained attorney versus assigned counsel/public defender? Does the type of representation or the individual prosecutor involved impact the pleas offered?

10. Far more People of Color are charged with and convicted of Resisting Arrest (ARO) than whites. Does this indicate bias in arrests by law enforcement and/or charging by prosecutors? Does this recognize the potential fear that People of Color may feel due to the disproportionate use of force their communities experience and witness? While that pulling away might be seen by an officer as resisting an arrest, it could also be a response to perceived danger. Or could this indicate that whites are pleading to lesser charges? What other charges are typically filed in these cases (such a Fleeing and Eluding, Attempted Resisting Arrest, or other)?

11. Felony Firearms (WFF) convictions carry a mandatory 2-year consecutive add-on sentence. Who gets charged or does not get charged with WFF? Specifically, when and for whom are WFF charges added on in AWIM, GBH or other cases? To what extent do WFF charges get dropped and for whom? How do these answers affect People of Color?

12. In weapons charges, Weapons-Concealed Carrying (WCC) is charged much more often than Felony Firearms or Firearms Possession, and there are few dismissals. Could this indicate that defendants plead to WCC to avoid more serious charges? Or that WCC is taken very seriously so dismissals are not offered? Could this also reflect lack of awareness of the license requirement? How do these answers affect People of Color?

13. What charges typically accompany armed robbery, assault with intent to commit murder, assault with intent to do great bodily harm less than murder, and homicide charges? Do they tend to differ depending on one's race? How do those added charges affect sentence outcomes between People of Color and white defendants?

14. As noted above, whites are charged with Assault with Intent to Commit Murder (AWIM) at a far lower rate than People of Color; even when whites are charged with AWIM, they do not

[54] https://www.sentencingproject.org/wp-content/uploads/2016/02/Facts-About-Prisons.pdf

RECEIVED by MSC 4/12/2022 1:07:49 PM

generally get convicted of that severe offense unlike People of Color. While AWIM conviction rates are lower overall than other FC cases we examined, the AWIM conviction rate for People of Color is still 46% higher than for whites.

In addition to the review of Judge Brown's criminal docket, a deeper review of court and prosecutor files on all assault cases should be undertaken to understand:

a. How and why the paths diverge when people of different race start out with a similar incident that results in the injury of another person.
b. The basis on which prosecutors charge assault cases differently.
c. Why People of Color (primarily Black people who make up 12.3% of the county's population) account for 80% of the AWIM charges.
d. Why AWIM conviction rates are overall lower than for other FC case categories yet still much higher for People of Color. What accounts for this wide disparity?
e. Whether prosecutors offer People of Color and whites similar plea deals and whether plea deals differ by race depending on the presiding judge.
f. Why the disparity in how some judges routinely convict People of Color of more serious offenses than whites.

15. Sentencing data for convictions for assault with intent to commit great bodily harm less than murder (GBH) show about 95% of People of Color are sentenced to prison and 5% to the less severe outcome of jail versus 62% of whites sentenced to prison and 37% to jail. What accounts for the more severe placement of People of Color in prison?

16. Of those sentenced to prison for a GBH conviction, the average minimum sentence for People of Color was more than a 64% higher than whites (when looking at all judges). The average minimum sentence for People of Color under one judge was an alarming 185.6% higher.

B. Further Analysis Using Other Sources

1. When a judge departs from the sentencing guidelines does that sentence reflect racial disparities? What data is available that the public can use to assess whether racial bias plays a role? Do the guidelines themselves reflect bias?

2. Are there ethical or other standards that Washtenaw County prosecutors follow in choosing particular charges for a case? Is there information available documenting the analysis prosecutors use in choosing charges? If there are standards, how is the public assured that prosecutors adhere to those standards/guidelines?

3. What processes are used in the prosecutor's office to monitor data related to possible racially disparate practices? How is that administered? What action is undertaken to correct racial disparities? How is it determined if a prosecutor exhibits racial bias or a practice within the office is racially biased? Is any racial disparity information and data available to the public? Who has responsibility for overseeing it?

RECEIVED by MSC 4/12/2022 1:07:49 PM

4. Is data available to document why prosecutors dismiss certain charges before or after the defendant reaches court? If so, does that reflect racial disparities in those decisions? Judges must approve plea deals, but do they have a role when charges are dismissed absent a plea deal?

5. Is there information that would explain where there is a small cohort of white defendants whether law enforcement arrested fewer whites, the prosecutor chose not to charge those that could have been charged or the prosecutor chose to charge them with a misdemeanor rather than a felony?

6. Black people represent 12.3% of the total population of Washtenaw County yet they account for 98.16% of the People of Color charged (FC cases) and 99.4% of the People of Color charged (FH cases) by prosecutors. What explains why People of Color are charged at exponentially higher rates than whites?

7. We suggest a deeper analysis of plea data to understand the disparities in pleas offered and charges dismissed.

8. How does a prosecutor's use of prior convictions impact the sentence s/he recommends to the court? What standards or procedures ensure that prosecutors apply prior convictions in a racially neutral way?

9. Is there information available that would allow us to determine which defendants were not designated as a Habitual Offender but could have been? How do these answers affect People of Color?

10. Prosecutors in Michigan have the ability to automatically charge children as adults without any hearing or deliberations before a judge. How often do the Washtenaw County prosecutors waive children into adult courts and is there a disparate impact on children of color?

11. For Washtenaw County offenders, do the terms of parole and probation follow the statutory requirements that they be related to the crime for which a person is on parole/probation? What percentage of returning Washtenaw County citizens violate their parole? For those deemed to have violated parole, what percentage of people are in violation for a substantive violation versus a minor violation? Do violations of parole/probation have a disparate impact on People of Color versus whites?

12. What discretion does a judge have in assigning weight to a pre-sentence investigation report (PSIR) at sentencing? Do they follow any standards? How is the public assured that a judge is not biased in the weight s/he gives to the PSIR?

13. Does any data and analysis exist by independent 3rd parties who have assessed any of the issues reflected in this report? Is that information publicly available? If not, why?

RECEIVED by MSC 4/12/2022 1:07:49 PM

14. For the categories we examined (other than homicide), we did not eliminate the higher charges. If we had the prosecutor's offense codes that rank the criminal violations in order of severity, further study of CREW's data could be done to eliminate the higher charges in each case category. Doing that would isolate the convictions in each case category to compare specific sentences for the same conviction.

15. How can we assess the impact of a particular judge on a prosecutor's decision regarding charges filed, whether the Habitual Offender designation is used against a defendant, whether and what charges to dismiss, whether a plea is offered, the prosecutor's sentencing request or other actions after the defendant is assigned to a particular judge? What role does the judge play in furthering the unequal treatment of People of Color by allowing prosecutors to use their discretion in a racially tinged manner?

16. Is there information available to show the process of assembling pre-sentencing investigation reports (PSIRs) and whether they are balanced for both sides? Does the close relationship between the probation staff at Michigan Department of Corrections (responsible for developing the PSIRs) and prosecutors impact fairness and transparency? Who decides what goes into the PSIR or is excluded? Is there a process for an individual to challenge a PSIR?

17. How does data regarding the issues in this report compare to other regions similar to Washtenaw County or to national data?

18. Should CREW or the proposed citizens racial equity commission consider inviting broader community input by community organizations, community members and returning citizens and their families impacted and involved in advocacy surrounding the criminal legal systems by launching listening sessions or an oral history project, such as a Story Corps Justice Lab, where people can share their experiences with the justice system?

VII. ACKNOWLEDGEMENTS

A. Biographical Information of CREW

Alma Wheeler Smith (Co-Chair) has been dedicated to public service for over 30 years and served in the Michigan Legislature for 14 years. Her legislative work earned her numerous awards from health, social, legal, and human services organizations. Governor Rick Snyder appointed Alma to fill a term on the Michigan Civil Rights Commission from February 2018 to December 31, 2019. She currently serves on the boards of the Regional Transit Authority of Southeast Michigan, the Michigan Campaign Finance Network, the Michigan Legislative Retirement System and is chair of the Board of Directors of University Bank of Ann Arbor. Alma lives in Salem Township (Washtenaw County).

RECEIVED by MSC 4/12/2022 1:07:49 PM

Linda Rexer (Co-Chair) was Executive Director of the Michigan State Bar Foundation for 30 years, leading the Foundation in its mission to provide funding and leadership for nonprofit legal aid for the poor. She chaired or served on numerous state level or national groups working to improve access to justice and has received state level and national awards for contributions to access to justice, the legal profession, and the public. Since her retirement, Rexer continues public service through positions on five nonprofit boards. She lives in Ann Arbor and is a lawyer admitted to practice in Michigan and before the U.S. Supreme Court.

Rev. Jerry Hatter has been the Pastor of Brown Chapel A.M.E. Church since 1991. For further information, please see https://www.bcamecy.org/pastor-hatter

Dan Korobkin is a civil rights and civil liberties attorney who has helped lead criminal law reform efforts in Michigan for over a decade. Korobkin's work and expertise spans a wide range of issues and includes debtors' prisons, bail reform, juvenile life without parole, police misconduct, and forfeiture. He lives in Ann Arbor.

Desiraé Simmons is a Co-Director with the Interfaith Council for Peace and Justice. She also is a community organizer and advocate serving in multiple grassroots organizations and coalitions. Currently she serves as the Vice Chair for the Michigan Democratic Party's Progressive Caucus and the Washtenaw County Democratic Party's Black Caucus. She is also a founding member of Liberate! Don't Incarcerate, Rising for Economic Democracy in Ypsi, What's Left Ypsi, and Untold Stories of Liberation and Love, among other groups. Simmons is a member of the Subcommittee on Housing Affordability and Accessibility through the City of Ypsilanti's Planning Commission.

Rev. Joe Summers has been the pastor of The Episcopal Church of the Incarnation since 1987. He and his congregation have been involved in criminal justice reform efforts throughout this period. Rev. Summers helped create and served as the county co-chair of the Washtenaw County branch of the Michigan Prisoner ReEntry Initiative and later became involved with Friends of Restorative Justice. He also serves on the Steering Teams of the Interfaith Council for Peace and Justice and the Washtenaw County Poor People's Campaign.

Organizational support and statistical analysis

MaryAnn Sarosi's legal career has focused on access to justice and civil rights issues as an executive director of a legal services program, assistant dean for public service at Michigan Law School and as a consultant. She currently sits on the Governor's Committee on Juvenile Justice.

Grady Bridges is an independent consultant, specializing in criminal justice research and policy. Prior to consulting, he worked for the Criminal Justice Policy Commission where his research focused on Michigan's sentencing guidelines, Raising the Age legislation, bail reform, and habitual offender sentencing.

RECEIVED by MSC 4/12/2022 1:07:49 PM

B. Acknowledgments

1. This report would not have been possible without the truly generous contributions of time and money from community members who supported this citizen-led effort. When folks heard about our work, they contacted us by the dozen and offered to roll up their sleeves and open their wallets. It made us realize we live in a compassionate, justice-minded community that wants to ensure that compassion and justice is applied equally to all of us.
2. One of the inspirations for this work was the Washtenaw County Board of Commissioner's *One Community: Advancing Racial Equity in Washtenaw County* initiative. We are confident that our report will greatly assist the Board in carrying out their commitment to advance racial equity in the county we call home.
3. Special thanks to CREW members for taking the leap of faith together and investing so much time analyzing and probing the data; to MaryAnn Sarosi whose idea sparked this project and persistence kept it going; and to Grady Bridges who exhibited skill and depth in helping us understand the data.
4. High fives and deep gratitude go to Gretchen Keppel-Aleks and Aaron Wolf for showing us how to leap into the data. We followed your lead and you made our work more productive.
5. There are so many case researchers to thank, including Rachel Manela, Ruthie Mills, Eleanor Mills, Gabriel Gurulé, Emma Huntley and Maria Kroeger.
6. Thanks to Phil Keller for donating his skills and time to developing CREW's website and to Steve Harrington whose deep expertise and creative problem-solving saved the day.
7. We are grateful for funding provided by the ACLU of Michigan which was enormously helpful in underwriting the cost of our effort.

RECEIVED by MSC 4/12/2022 1:07:49 PM

VIII. TABLES & METHODOLOGY MEMO

A. Acronyms

1. Non-capital felony (FH)
2. Capital Felony (FC)
3. Assault, Resisting or Obstructing an Officer (ARO)
4. Controlled Substance-Delivery (DEL)
5. Controlled Substance-Possession (POS)
6. Suspended License (SLI)
7. Weapons-Carrying Concealed (WCC)
8. Weapons-Felony Firearm (WFF)
9. Weapons-Possession by Felon (WPF)
10. Armed Robbery (ARM)
11. Assault with Intent to Commit Murder (AWIM)
12. Assault with Intent to Do Great Bodily Harm Less than Murder (GBH)
13. Homicide (HOM)

B. Data Tables

C. Statistician's Methodology Memo

TO: Citizens for Racial Equity in Washtenaw (CREW)
FROM: Grady Bridges, Consultant
DATE: October 16, 2020
SUBJECT: Data and Methodology for CREW's Analysis

In the recent report, "*Race to Justice"*[55], CREW examined charging and sentencing practices in the 22nd Circuit Court in Washtenaw County. Using data gathered from the court's website, they analyzed over 1,000 felony cases and identified several concerning disparities between people of color (POC) and white defendants. This memo presents an overview of the dataset constructed by CREW, the methodology used to analyze sentencing outcomes, and additional data sources.

Data Collection

Utilizing the circuit court's website[56], CREW collected data on all noncapital felony cases (FH cases) filed in Washtenaw County between 2017-2019 and all capital felony cases (FC cases) filed in

[55] *"Race to Justice - Citizens for Racial Equity In Washtenaw's Report On Racial Disparities in the Washtenaw County Criminal Legal System"*, https://www.citizensforracialequitywashtenaw.org/crew-s-report
[56] Criminal Case Records, https://tcweb.ewashtenaw.org/PublicAccess/default.aspx

RECEIVED by MSC 4/12/2022 1:07:49 PM

Washtenaw County between 2013-2019. While data was compiled on all cases filed during these periods, not every case could be included in the analysis. The following cases were excluded from the analysis:

Exclusion	Rationale
• Cases where the defendant's race was unknown	Cannot compare outcomes by race for these cases
• Cases that are currently pending	Outcomes for charges in these cases are unknown
• Cases in which the defendant was not convicted of any charge	These cases would cause different samples for the charging and sentencing analyses

Crime Categories

Unfortunately, the publicly available data regarding the offenses charged in each case was limited. While general description of crimes and some broad statutes are included in the online records, neither allows for easily identifying the same criminal charges across cases. Due to this limitation, conducting a broad analysis comparing outcomes across all crimes was not a practical option. As an alternative, CREW identified a subset of crimes to examine, selecting all cases that included any charges for the 11 categories listed below. In total CREW reviewed the charging and sentencing outcomes for 1,014 cases (265 FC cases, 749 FH cases).

- **FC Crime Categories**: Armed Robbery (ARM), Assault with Intent to Commit Murder (AWIM), Assault with Intent to Commit Great Bodily Harm Less than Murder (GBH), Homicide (HOM)
- **FH Crime Categories:** Assault, Resisting, Obstructing an Officer (ARO), Controlled Substance-Delivery (DEL), Controlled Substance-Possession (POS), Suspended License (SLI), Weapon-Carrying Concealed (WCC), Weapon-Felony Firearm (WFF), Weapon-Possession by Felon (WPF)

Case Consolidation

While reviewing the case data we identified individuals that had multiple cases filed against them. Given our dataset spans several years it is possible that defendants may recidivate, resulting in additional cases. However, several of these cases were filed within a short time of one another and were sentenced on the same date, suggesting that the cases were consolidated by the court.

Distinguishing which cases fall under the latter scenario is necessary for producing accurate case statistics for POC and white defendants. Failing to account for these consolidations could lead to double-counting the same case and improperly weighting the summary statistics. For example, consider two cases filed against a defendant in the same week that are consolidated by the court and disposed at the same time. Furthermore, suppose the defendant was a person of color and received a prison sentence of 5-10 years. Ignoring the consolidation and treating these as separate cases, would incorrectly count the same defendant twice, adding an additional 5-10 year sentence to the average of POC defendants.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Given the potential impact from ignoring this scenario, we considered multiple FH cases filled within the same week as one case. For FC cases, multiple cases filed within 60 days were consolidated into the same FC case. A more conservative approach was chosen for the FC cases to capture the extended judicial process associated with such serious cases. Taking this approach, the number of FC cases was reduced from 34 FC cases were consolidated reducing the number FC cases from 283 to 265 cases. Similarly, 172 FH cases were consolidated reducing the number of FH cases from 850 to 749 cases.

Methodology

The tables[57] in CREW's report address two different sets of questions regarding outcomes for defendants in Washtenaw County:

- Question 1 (Teal Tables): For defendants ***charged*** with a particular crime, are there significant differences in charging practices between POC and white defendants (e.g., how often are plea deals and habitual offender enhancements used)?
- Question 2 (Orange Tables): For defendants ***convicted*** of a particular crime, are there significant differences in sentencing outcomes, such as the type of sentence (i.e., probation, jail, or prison) and the sentence length?

An important distinction exists between the cases used to answer the first question and those used for the second. The teal tables include cases where the defendant was charged with the specific crime, while the orange tables include cases where the defendant was charged and convicted of the specified crime.

To compare prosecution practices, cases charged but not convicted of the specific crime still need to be considered. This is because these cases can be the result of plea deals to reduce or drop charges. Plea deals, due to their discretionary nature, are an important factor when identifying disparities between POC and white defendants.

Conversely, cases without a conviction for the specific crime are not suitable when comparing sentencing outcomes. Including such cases would conflate the disparities from prosecution with the sentencing decisions. Suppose prosecutors reduce AWIM charges to GBH more often for white defendants. As a result of these plea deals, the judge sentences the white defendants for GBH. Including these cases in the AWIM sentencing analysis, would overestimate the disparity in sentence length between POC and white defendants. For this reason, all cases included the sentencing tables were convicted of the specific crime.

To address the second question, the average sentence lengths are compared between POC and white defendants. As noted in footnote 7 of the *"Race to Justice"* report (p. 4), these averages incorporate concurrent and consecutive sentences for additional convictions in the same case. This is necessary to capture the actual type of sentence imposed as well as the total sentence imposed. For example, suppose a defendant was convicted of two different felonies and sentenced to concurrent terms. For the first conviction they were sentenced to 9 months in jail, while the second conviction resulted in a 1-2 year prison sentence. In this example the second conviction determines the type and length of sentence for this case. Ultimately, the 9-month sentence for the first conviction would be served in prison concurrent

[57] FH and FC Data Table, https://www.citizensforracialequitywashtenaw.org/data-tables

RECEIVED by MSC 4/12/2022 1:07:49 PM

to the 1-2 year sentence. Furthermore, the actual length of time before the defendant would be eligible for release would be 1 year, not 9 months.

While this approach accurately captures the type of sentence imposed, limitations can arise in scenarios where the defendant was convicted of an additional felony that was more serious that the specific crime in each table. Under these conditions the total sentence length would reflect the more serious conviction. The impact of this limitation would be most noticeable in cases where the more serious conviction carries a drastically longer sentence, for example FC cases with convictions of GBH. Apart from the GBH cases, the impact from this issue is likely minimal, due to the seriousness of the other crime categories selected by CREW. For example, it is unlikely that an individual convicted of AWIM would have an addition conviction with a much longer sentence. In the rare circumstances where this could occur the other sentence would likely have to be life imprisonment, which would remove the biased case from average sentence calculations anyway.

Most importantly, these scenarios alone cannot account for the differences between POC and white defendants. The likelihood of having additional and more serious convictions would need to be correlated with the defendant's race for the results to be biased. Without this correlation, the effect of this limitation would be shared by both POC and white defendants.

One alternative to the approach taken in the report would be to compare the sentences for just the specific crime category and ignore concurrent and consecutive sentences. Taking this approach would address the issue caused by more serious charges, however it is not without limitations. In particular, the type of sentence for an individual conviction may not reflect the actual sentence type for that case. Reporting outcomes in this manner may not accurately reflect the actual number of defendants that receive sentence for probation only, jail, and prison.

Additional Data

CREW's report has highlighted serious issues within Washtenaw county's criminal justice system. Considering these findings, it is vital that efforts are made to continue this review and expand the scope of the analysis. Unfortunately, doing so would require additional information that is not made publicly available by the court. One source of crucial information that is missing are presentence investigation (PSI) reports.

PSI reports are produced by probation officers for all felony sentencings and provide detailed information about the case and the defendant. This includes the most serious offense in a case, the defendant's prior criminal record, and a recommended range for the sentence. As noted in CREW's report (p. 47), this information could be used to identify the severity of charges in each case, thus removing the limitations discussed above. Furthermore, if the court were to make the data from these reports available, CREW could expand their analysis to:

- Account for the impact of the defendant's prior record and habitual status on sentencing decisions.

RECEIVED by MSC 4/12/2022 1:07:49 PM

- Compare outcomes between POC and white defendants based on the sentencing guidelines recommendations, expanding the analysis far beyond the 11 crime categories examined in the report.
- Determine how the often the sentences for POC and white defendants depart from the recommended range provided by the sentencing guidelines.

RECEIVED by MSC 4/12/2022 1:07:49 PM

APPENDIX I

Resentencing Transcript

RECEIVED by MSC 4/12/2022 1:07:49 PM

STATE OF MICHIGAN

THIRD JUDICIAL CIRCUIT COURT - CRIMINAL DIVISION

THE PEOPLE OF THE STATE OF MICHIGAN

-vs-

DAMETRIUS BENJAMIN POSEY, Case No. 18-000074-01

Defendant.

- -/

COURT REPORTING SERVICE
CRIMINAL DIVISION-THIRD
2020 FEB 14 AM 9:51

RESENTENCING

Proceedings had and testimony taken before the Honorable ULYSSES W. BOYKIN Judge of the Third Judicial Circuit Court, Criminal Division, at Room 603, Frank Murphy Hall of Justice, Detroit, Michigan, on Thursday, November 7, 2019, commencing at or about the hour of 10:40 A.M.

APPEARANCES:

STEVE ADAMS
Assistant Prosecuting Attorney
Appearing on behalf of the
People of the State of Michigan

ADRIENNE N. YOUNG
Appearing on behalf of Defendant

Regenia S. Veasy, C.S.R. (R-2350)

RECEIVED by MSC 4/12/2022 1:07:49 PM

Detroit, Michigan

Thursday, November 7, 2019

About 10:40 A.M

--- --- ---

THE CLERK: This is Case Number 18-00074, People versus Dametrius Posey. This matter is before the Court today for resentencing.

THE COURT: Good morning.

May we have appearances, please, for the record?

MR. ADAMS: Good morning, Your Honor -- Steve Adams, P 37724, for the People.

MS. YOUNG: Good morning, Your Honor -- Adrienne N. Young, from the State Appellate Defender Office, P 77803. I'm here on behalf of Mr. Posey who joins me at counsel table.

THE COURT: All right.

Have you had an opportunity to review the Pre-sentence Investigation Report?

MS. YOUNG: I have.

THE COURT: Are there any additions, corrections, deletions, or modifications required?

MS. YOUNG: Yes.

Mr. Adams and I have agreed to some modifications. I'll state them for the record, and then, Mr. Adams, please feel to jump in if you feel I've stated something

RECEIVED by MSC 4/12/2022 1:07:49 PM

in error.

Page 1 of the new Pre-sentence Investigation Report, begins with Charge 15, and then goes to Charge 16, which is assault with intent to commit great bodily harm less than murder. That should actually read assault with intent to commit murder, the same as the first charge. There should be two AWIM charges.

THE COURT: So, Charge 16, which is the second one listed on the face-sheet should be assault with intent to commit murder?

MS. YOUNG: That's correct. It should duplicate Charge 15 above it.

THE COURT: And, that would have the maximum of life. All right, I'm making that correction.

All right?

MS. YOUNG: One of the felony firearm charges we'll just say the last one on the table for ease in editing should just be crossed off. And, the reason being in addition to the felony firearm charges that we agreed to remove, one more has come to our attention. It was filed in connection with a C.C.W. charge, which is against the statute.

THE COURT: C.C.W. is not a predicate felony of felony firearm.

MS. YOUNG: Right. So, for that reason we would

RECEIVED by MSC 4/12/2022 1:07:49 PM

remove any additional felony firearm that's listed.

THE COURT: So, we're going to remove Charge 28?

MS. YOUNG: I have that one crossed out, again, just for ease of editing. I don't know if that's necessarily as directed.

Mr. Adams, do you have a preference as to which felony firearm charge is removed?

THE COURT: Well, I want to make sure we remove the right one.

MR. ADAMS: Your Honor, if you want to remove the right one -- and, I think we need to number the Counts correctly. For example, there is no Charge 15. This Defendant was charged with Charge Number 1, and then the next one down wouldn't be 16, but it would be 3.

THE COURT: All right.

MR. ADAMS: And, so on. I don't know if that's important to make those changes on the Pre. I know it's correct on the Judgment Of Sentence, at least, the one that was filed earlier.

THE COURT: Let's have them so they're in conformity.

MR. ADAMS: Thank you, Your Honor.

THE COURT: All right.

Charge Number 15 should be Charge Number 1?

RECEIVED by MSC 4/12/2022 1:07:49 PM

MR. ADAMS: Yes, Your Honor.

Charge Number 16 should be Charge Number 3.

THE COURT: All right. Charge 16 should be 3.

MR. ADAMS: Charge Number 19 should really be Number 5.

THE COURT: Charge 19 should be Number 5.

THE CLERK: Judge, this is what I have. I don't know if that will help you, but I printed it.

THE COURT: Okay. Let's see.

Charge 1 is assault with intent to murder. Charge 3 is assault with intent to murder. Charge 5 is carrying a concealed weapon. Charge -- well, let's see -- the one that's 21, is that Charge 6?

MR. ADAMS: That's 7, Your Honor. Charge 6 is the one that's 20, carrying a weapon with unlawful intent.

THE COURT: Okay. So, then the one that charged 21 should be?

MR. ADAMS: Charge 7.

THE CLERK: And, what's the charge?

MR. ADAMS: The charge there is felon in possession, Your Honor.

THE COURT: Okay, felon in possession. All right?

MR. ADAMS: And, then 8 --

RECEIVED by MSC 4/12/2022 1:07:49 PM

THE COURT: (Interposing) Charge 24 it should be 8?

MR. ADAMS: No. No. I'm sorry. That's not right. There isn't an 8. Charge 8 was only Defendant Number 2.

THE COURT: Okay.

MR. ADAMS: The next one is Number 9, and that's the felony firearm for the AWIM. So, that's listed here as 24, so 24 should be 9.

THE COURT: Charge 24 should be Charge 9, and that's the felony firearm for the AWIM?

MR. ADAMS: Yes, Your Honor.

THE COURT: All right. Let me put that in.

MR. ADAMS: Now, the next one up is Charge 25, and that should be 11. That's another felony firearm for the other AWIM for the second victim.

THE COURT: Now, 10 for this Defendant?

MR. ADAMS: Charge 10 was a G.B.H. That conviction, along with some others are being vacated.

THE COURT: Yes, okay.

MR. ADAMS: But, it was with this Defendant.

THE COURT: All right.

So, Charge 25 is now Charge 11, and that's for felony firearm also for the AWIM?

MR. ADAMS: Yes, Your Honor.

RECEIVED by MSC 4/12/2022 1:07:49 PM

THE COURT: All right?

MR. ADAMS: And then 13 and 14 -- I'm sorry.

THE COURT: So, 26 --

MR. ADAMS: (Interposing) Excuse me just a second -- I believe that's 10. There is no 8.

MS. YOUNG: No. And, 10 was 1 that got dismissed.

MR. ADAMS: Okay. So, 25 is 11.

MS. YOUNG: Right.

MR. ADAMS: I'm sorry.

THE COURT: And that's for felony firearm for the AWIM.

MR. ADAMS: Yes, Your Honor. That's it. I'm sorry. I lost track.

So, what's listed here as 26 would be a felony firearm for the carrying a weapon with unlawful intent. And that Count XIII.

THE COURT: So, 26 is 13.

MR. ADAMS: And the last is Number 27, which should be 15.

THE COURT: So, 27 is 15.

MR. ADAMS: And, that's the felony firearm for felon in possession.

THE COURT: All right.

MR. ADAMS: And, as Counsel stated, the last one that's listed here as 28 should be eliminated because

RECEIVED by MSC 4/12/2022 1:07:49 PM

that would correspond with the felony firearm for the C.C.W.

So, that's a very longwinded way of saying yes, I agree with Counsel. The last one is the one that should be crossed off.

THE COURT: All right.

MS. YOUNG: That would be Charge 14 gets crossed off, is that right?

MR. ADAMS: Well, I have it as -- no. Actually it was 13 on the Information.

MS. YOUNG: That's right. Because right now we have Charge 26 as 13, which we said is associated with the unlawful intent one.

MR. ADAMS: The unlawful intent one on the Information is 14.

MS. YOUNG: Okay.

So, Judge, Charge 26 which reads -- you have it as 13 now?

THE COURT: Yes.

MS. YOUNG: It should read 14.

MR. ADAMS: Right.

THE COURT: All right.

MR. ADAMS: And, then 27 is 15, right?

MS. YOUNG: Yes. And, 27 is still 15. And, 28 which drops off is 13.

RECEIVED by MSC 4/12/2022 1:07:49 PM

THE COURT: All right.

MS. YOUNG: Thank you, Mr. Adams.

THE COURT: All right.

Are there any other corrections?

MS. YOUNG: Related to the ones we've made on this table, where it says plea agreement, and it's a summary of the agreed upon changes, but rather than dismiss Count III, the first number should be Count II.

THE COURT: Is Count III still in there?

MS. YOUNG: It is. Count III is what was listed as Charge 16 so -- oh, yes. We deleted III and changed it to II. So, III still stands, and II was dismissed pursuant to an agreement.

THE COURT: The plea agreement, dismiss Counts II --

MS. YOUNG: (Interposing) And, IV, X, XII, and then we'll add XIII, which was the change we made to drop off the bottom of the table.

THE COURT: All right.

MS. YOUNG: That's all the changes I have for the first page.

Mr. Adams, anything else?

MR. ADAMS: I have nothing further.

MS. YOUNG: Okay.

And then the second page --

RECEIVED by MSC 4/12/2022 1:07:49 PM

THE COURT: Of the cover-sheet?

MS. YOUNG: Yes.

THE COURT: All right.

MS. YOUNG: This will be on the jail credit table.

THE COURT: All right.

MS. YOUNG: Again, with agreement with Mr. Adams and his office, the total days jail credit currently reflects zero. It should be 436. The reason it reflects zero is because Mr. Posey was on parole at the time of this offense, but since being sentenced for this offense he's served 436 days. So, while it remains zero for his pretrial incarceration, he does get credit for the over year's time he's spent in MDOC. custody. So, that total is 436 days.

THE COURT: So, is that the same dates that are covered?

MS. YOUNG: So, the new date would be a start date. So, if we keep 10/8/17, and rather than go to the present date, we would go to August 27 of '18.

THE COURT: So, 8/27/18?

MS. YOUNG: Um-hum.

THE COURT: Yes?

MS. YOUNG: And, we make a new --

THE COURT: (Interposing) Is that a yes?

RECEIVED by MSC 4/12/2022 1:07:49 PM

MS. YOUNG: Yes.

THE COURT: Um-hum cannot be picked up by my reporter.

MS. YOUNG: Yes, Your Honor.

THE COURT: Okay.

MS. YOUNG: So, now the line on the table says 10/8/17 to 8/27/18, Action, Arrest, Sentence Details, WCJ. I don't believe there's any MDOC custody there, so we can cross out MDOC and just leave WCJ.

THE COURT: And, the total jail credit would be again?

MS. YOUNG: I'm going to do the math right now.

THE COURT: It's 324 days.

MR. ADAMS: Yes, Judge. Those are the days for which he does not get any credit.

MS. YOUNG: Correct.

THE COURT: So then --

MR. ADAMS: (Interposing) Because then there would be another line to that credit.

THE COURT: All right.

So, instead of the 760, it's 324.

MS. YOUNG: That's correct, Your Honor.

THE COURT: But, then the total jail credit is zero.

MS. YOUNG: That's correct, Your Honor.

RECEIVED by MSC 4/12/2022 1:07:49 PM

THE COURT: All right.

MS. YOUNG: Then we'll have a new line on the table, or a new table. I don't know how best to visualize this, but it would be from 8/28/18 to the present date, 11/7/19.

THE COURT: That's 8/28/18 to?

MS. YOUNG: To 11/7/19. Action would be MDOC. Sentence details would also be MDOC.

THE COURT: And, the total number of days.

MS. YOUNG: It's 436.

THE COURT: And, then that 436 on the first face-sheet would have to be reflected on the -- for the felony firearm days jail credit.

MS. YOUNG: That's correct, Your Honor.

THE COURT: All right.

MS. YOUNG: And that concludes the edits for page 2.

THE COURT: All right.

MS. YOUNG: I believe that's correct, right, Mr. Adams.

MR. ADAMS: I don't have any quarrel with that.

THE COURT: All right.

MS. YOUNG: The next is the Evaluation and Plan page. And I just have one requested added to that. It would be the paragraph that begins to towards the bottom

RECEIVED by MSC 4/12/2022 1:07:49 PM

of the page, the Defendant is unemployed.

THE COURT: So, that would be the fifth paragraph from the top.

MS. YOUNG: The fifth paragraph, third sentence -- The Defendant has a juvenile record. It goes on to say that the one misdemeanor was warned and dismissed.

THE COURT: Yes.

MS. YOUNG: So, I'd like that entire sentence deleted because I believe it's inaccurate to say that Mr. Posey has a juvenile regard.

MR. ADAMS: Well, I'd like to be heard on that.

THE COURT: All right t.

MR. ADAMS: Your Honor, if you look at the possible dispositions in Juvenile Court that follow an adjudications, warn and dismiss is one of them. It doesn't mean that the charge was dismissed. It just means that as disposition, it's like a sentence. There was nothing that was required by the Court following the disposition. This is contrast, for example, to an juvenile history on page 4 and 5, the B and E building charge which was charge which was flat out dismissed, as it says Notes: case dismissed on 6/19/01.

Warrant Number 3 of 3, which was flat out dismissed on 7/1/05, without any adjudication. What record means to me is adjudication, and it does

RECEIVED by MSC 4/12/2022 1:07:49 PM

accurately state the disposition of warned and dismissed, but it's still a record.

THE COURT: Anything further on that, Miss Young?

MS. YOUNG: No, Your Honor.

THE COURT: I would agree with the People's assessment. Warning is a disposition. Now, what was done and whether it was just a flat dismissal it's factually accurate, so it will remain under Evaluation and Plan, in the fifth paragraph on page 1.

MS. YOUNG: Thank you for your consideration, Your Honor.

THE COURT: All right. Anything else?

MS. YOUNG: I have nothing further.

THE COURT: Anything on behalf of the People, corrections or modifications, or deletions?

MR. ADAMS: In terms of the body of the report, no. I do have some guideline issues, but nothing in terms of the report.

THE COURT: All right. Let's go to the guidelines then.

All right, on the guidelines?

MR. ADAMS: Your Honor, I believe that OV 6 is mis-scored. OV 6 measures the Defendant's intent at the time of the incident, and it's only charged in assault with intent to murder cases and homicide cases.

RECEIVED by MSC 4/12/2022 1:07:49 PM

Then OV 6 -- and the rule is that the Court is required to score it consistently with the jury verdict -- and, of course, here the jury verdict was two Counts of assault with intent to murder -- unless there is information that was not presented to the jury. I'm not aware of any information that was not presented to the jury one way or the other, but I am aware of the information in the Pre-sentence Report which Counsel did not challenge on page 3, the third line which indicated that both Defendants produced handguns and attempted to rob the Complainants.

Now, OV 6 is scored fifty points when the offender had a premeditated intent to kill or the killing was committed while committing or attempting to commit a robbery. Then there is other crimes, too, the other one to go along with felony murder. And, of course, there it says the killing was committed, and there was no killing in this case. And, perhaps, this doesn't apply, but in looking at this, I did see in a slightly different context the Court of Appeals affirmed an publish opinion -- excuse me -- the scoring of fifty points for assault with intent to murder conviction, however, that was when the aggravating factor was not a felony like robbery, but when the aggravating factor was premeditation. And, that's a

RECEIVED by MSC 4/12/2022 1:07:49 PM

little bit more consistent with the language that the legislature used.

Alternatively, my first argument is it's fifty points given that the Pre-sentence Report indicates that there's a robbery that was being attempted. Thus, alternatively, Your Honor, I don't see any way that could avoid at least twenty-five points for unpremeditated intent to kill given the jury's verdict.

THE COURT: All right. Let me hear Defense Counsel?

MS. YOUNG: So, a couple of things, Your Honor.

First, the case that Mr. Adams refers to, People versus Steanhouse, is a drastically different set of facts where in part the Defendant watched the person bleed out and showed greater signs of intent.

But, more importantly, what I would refer this Court to is the original sentencing transcript at page 7 where the Court described this offense as an extreme emotional state cause by adequate provocation before a reasonable amount of time elapsed for the offender to calm. There is gross negligence amounting to an unreasonable disregard for life.

That definition basically fits a manslaughter type of offense. That description of this offense, which is reflected in the fact that there was an intent to rob,

RECEIVED by MSC 4/12/2022 1:07:49 PM

and then this court seen the video, and there's complete melee where one of the victims was actually the primary shooter. Most of the bullets were from his gun.

Given the emotional state that this court observed at the original sentencing, I think that precludes the scoring of fifty points for OV 6.

With respect to the current scoring, I think that the probation department chose to score OV 10 this time, and at the previous hearing because of that emotional melee this court observed itself, and the sentencing transcript on page 7. I think there was a very chaotic situation wherein there wasn't even an unpremeditated intent to kill, but rather just a chaotic result wherein the victim was an active shooter as well, albeit one who was lawfully carrying a gun. So, Your Honor, I would ask that this court maintain the scoring the probation department has now stuck with twice which is ten points, but certainly I believe this court should be precluded from fifty points.

THE COURT: Let me hear from the People again?

MR. ADAMS: Thank you.

Your Honor, if there really was a melee, if the killing would have been manslaughter -- I mean if the shooting would have resulted in a manslaughter had the Complainant died, then assault with intent to murder

RECEIVED by MSC 4/12/2022 1:07:49 PM

would have been an illegal verdict.

THE COURT: That's correct. You have to have the specific intent to kill.

MR. ADAMS: Thank you.

And, since the instruction is unless the sentencing court has information that was not presented to the jury, an OV 6 score must be consistent with the jury verdict, it would seem to me that at the very least twenty-five points are called for.

There was no jury verdict on the question of the robbery. He wasn't charged with assault with intent to rob, and I don't have that argument with respect to that. My argument is simply based on the fact that that's the way it's characterized in the Pre-sentence Report, which has not been challenge. But, that would get fifty points if the Court agrees.

THE COURT: All right.

In looking at -- this was a jury case. The jury did find the Defendant guilty of assault with intent to murder. There has to be, in the third element, that the Defendant intended to kill the person he had assaulted and the circumstances did not legally excuse or reduce the crime. It's a specific intent crime.

And, in People versus Taylor, 422 Mich 554, it

RECEIVED by MSC 4/12/2022 1:07:49 PM

says -- assault with intent to murder requires an actual intent to kill, an intent to inflict great bodily harm or a wanton or wilful disregard of the likelihood that the natural tendency is to cause death or great bodily harm, while sufficient to support a conviction of second degree murder is insufficient to support the conviction of assault with intent to murder. A specific intent to kill is the only form of malice that supports the conviction for this offense -- People versus Cochran, 155 Mich Ap 191 --

MS. YOUNG: (Interposing) Your Honor, may I?

THE COURT: Yes.

MS. YOUNG: I would just point out that at no point in the case law that you've referenced, or in the arguments thus far that there have been any indication of the premeditation. So, for example, the twenty-five point score which this Court agreed to the last time affects an unpremeditated intent to kill.

In the case law that you're reading, the intent to murder has to be there. But, murder comes in various forms, in the case law of this state, and then the statutory law of this state, one of them being unpremeditated or done in an emotional state, and that's exactly what this court found given the exact same convictions it was faced with including AWIM at the time

RECEIVED by MSC 4/12/2022 1:07:49 PM

of sentencing.

So, I would ask that, at the very least, this court sticks to its original factual findings given that no additional evidence has been presented of the circumstances to make the situation any different. So, given that AWIM does not require the premeditation that was fifty points scored by this Honorable Court --

THE COURT: (Interposing) How does that square with People versus Taylor, 422 Mich 554, which says that the elements of second degree murder would not be sufficient to support an assault with intent to murder?

MS. YOUNG: I understand that.

MR. ADAMS: Judge, I think that the difference between fifty and twenty-five is in our case whether this was done while committing or attempting to commit a robbery. If it was, then fifty points is the appropriate score by right reading of the rule, and there's an argument that you would have to have a killing in order to get that score. But, that was not the case in this Steanhouse case that I alluded to and Counsel mentioned by name. The --

MS. YOUNG: (Interposing) Your Honor, I just want the record to be clear that --

THE COURT: (Interposing) Wait a minute. Let him finish.

RECEIVED by MSC 4/12/2022 1:07:49 PM

MS. YOUNG: Okay.

THE COURT: Were you finished?

MR. ADAMS: I'm not saying -- just to say one more thing.

I'm not saying that there's evidence of premeditation in this case. I don't know anything about whether there was premeditation in this case or not. I only know that there was an intent to rob because it was in the Pre-sentence Report.

MS. YOUNG: And, the plain language of the statute says that a killing must occur with, and then a list of felonies, during an intent to rob. The only time where a killing is not required under the statute for fifty points is if there is a premeditation proven. Here, Counsel has conceded that there isn't evidence of premeditation which is in line with this court's findings at page 7 of the original sentencing. Only if there is a killing, which there is not here, is it relevant that this occurred during a robbery. The plain language of the statute requires a killing done during the course of a robbery. That's expressed. So, again, I'd argue that this court cannot score OV 6 at fifty points. I'd be open to discussion about twenty-five points.

THE COURT: It was scored at twenty-five

RECEIVED by MSC 4/12/2022 1:07:49 PM

previously?

MS. YOUNG: Yes, Your Honor.

THE COURT: I'll score it at twenty-five now then. So, OV 6 will be scored at twenty-five points.

MR. ADAMS: Thank you, Your Honor.

I have no other quarrels with the sentencing guidelines.

MS. YOUNG: Just for additional --

THE COURT: (Interposing) Hold on just a second.

MS. YOUNG: Okay.

THE COURT: So, OV 6 will now be scored at twenty-five points instead of twenty.

MS. YOUNG: That's correct. That would take the total OVs to --

THE COURT: (Interposing) Anything else under the offense variables, either Counsel?

MS. YOUNG: I have nothing else at this time.

THE COURT: The People have nothing else on the offense variables.

MR. ADAMS: Correct, Your Honor.

THE COURT: So, that's an additional fifteen points, which would take that up to eight-five points?

MR. ADAMS: Yes, Your Honor.

MS. YOUNG: That would bring the OV level which

RECEIVED by MSC 4/12/2022 1:07:49 PM

is listed as IV to V.

THE COURT: And, this is on the A Grid, is it?

MS. YOUNG: Um-hum. And, so I have the --

THE COURT: (Interposing) Yes?

MR. ADAMS: Yes, Your Honor.

MS. YOUNG: Yes, Your Honor.

THE COURT: Okay,

MS. YOUNG: I have it as being 171 to 427, A Grid, Hab 3.

THE COURT: It's E-V on the A Grid, Hab 3, so that's 171 months to 427 months, okay?

MS. YOUNG: Yes, Your Honor.

MR. ADAMS: Yes, Your Honor.

THE COURT: And, that's under the offense variable?

MR. ADAMS: Yes, Your Honor.

MS. YOUNG: Yes, Your Honor.

THE COURT: All right.

Now, back to the Defense, anything you want to say on behalf of Mr. Posey with respect to sentencing?

MS. YOUNG: Yes, Your Honor.

Just some brief comments from me, and then Mr. Posey has also prepared comments for the Court today.

THE COURT: He'll get his chance after the People have entered their comments.

RECEIVED by MSC 4/12/2022 1:07:49 PM

MS. YOUNG: Thank you, Your Honor.

In addition to Mr. Posey's limited juvenile record that this court has already -- that we've already discussed today in court, Mr. Posey only has two prior felonies from 2005, so they are arising out of the same incident. And, importantly, his growth since that time, and since his incarceration for the present offense, falls right in line with MDOC and social science policies.

So, by that I mean, what I pointed to in my sentencing memo is that the MDOC has a policy wherein they score inmates a particular number of points until they're twenty-six. And, after twenty-six, there's expected to be a drastic change in their emotional and behavioral development. Mr. Posey reflects that change. He reflects why that policy is in place.

He's thirty-one years old at this time, has obtained his G.E.D. As this court can see from his MDOC record, he's on the waiting list for voluntary classes, not required ones, maintains a strong record there, and importantly, that sort of mental behavioral development that social science and the MDOC predicts will happen, it's really borne out in the letters of support from Mr. Posey's family. This court has received between the original sentencing and now at

RECEIVED by MSC 4/12/2022 1:07:49 PM

6

least twenty letters of support on Mr. Posey's behalf. All of them reflect more than just a base knowledge or a base desire to express their love for a family member, but also reflects this growth that I'm talking about, this development, this accountability that Mr. Posey has for his actions. And, I've provided several particularly poignant quotes in bold and the sentencing memorandum that I submitted to this court. Among them which includes from a sister -- he's made mistakes, and takes full accountability for his actions.

THE COURT: Is that on page 4 of your memorandum?

MS. YOUNG: Yes.

THE COURT: Under paragraph 9?

MS. YOUNG: Yes, Your Honor.

THE COURT: Okay.

MS. YOUNG: And, it also reflects, and on that same page and other pages, not only is Mr. Posey receiving a great degree of family support, but he's also found God again. He has a support system that's not reliant just on those family members, but that also he's actively in pursuit of his faith and of falling in line with principles that he reads in the Bible.

Even in my visit to him just two days ago, he found himself quoting from the Bible in preparation for this hearing today. So, he not only has a strong family

RECEIVED by MSC 4/12/2022 1:07:49 PM

support system, but he has a strong faith in God that will guide him in the right direction from here forward, and continues to while he's in MDOC custody.

As this court knows from its reading of the sentencing memorandum, I recognize first at page 3, just sort of the math piece involved. So, originally with the guidelines, this court sentenced him to eleven percent higher than the bottom of the guidelines range. So, I did that math for this court, did the correctly scored guideline range. That would be 189 months. That's eleven percent of the newly scored guidelines. So, that we think would be exactly what this court originally sentenced.

So, now this court has more information. It has new letters from family, it has new letters from community members, and it has an MDOC record albeit a short one. So, with this additional information, I'd ask this court to consider a sentence of 120 months.

Yes, this would be a downward sentence from the guidelines scored, but I think it would be the just sentence. It would be one that would deter others from committing similar acts, but it would be just given the individual Mr. Posey has become, and the person he shows the potential of becoming.

I note in my pleading that punishment that makes

RECEIVED by MSC 4/12/2022 1:07:49 PM

it harder for people to return to society as law-abiding productive citizens should be avoided. Any sentence longer than ten years, especially given that we have him convicted as felony firearm, so it would really be fifteen -- should be avoided. And, I believe that the just result would be for AWIM, 120 months sentence. Thank you, Your Honor.

THE COURT: All right. Let me hear from the People.

MR. ADAMS: Thank you, Your Honor.

Your Honor, I didn't handle the trial, of course, or have access to trial transcripts, me personally, so I'm just going to make two observations.

First, assuming that Counsel and the Department of Corrections are correct, and one would expect a growth, I don't know exactly the word she used, but it's significant improvement in a person's cognition and behavior at age twenty-five, I would point out Mr. Posey turned twenty-five on June 25, 2013. And, he was returned to prison as a parole violator first on October 14, 2013, he was paroled again in April of 2015, and returned as a violator, a technical violator on January 27, 2016. This is all after he's had this change of brain that Counsel and the Department of Corrections references. He's out for about a year on parole for

RECEIVED by MSC 4/12/2022 1:07:49 PM

assault with intent to murder and felony firearm when he commits two more assaults with intent to murder and felony firearm second.

The guidelines -- the sentence that the Court gave him initially was within the guidelines as they are now calculated. I submit that the behavior that's been exhibited during the brief time -- the relatively brief time he's been in prison isn't sufficient to warrant a significant change or to ask the Court not to reaffirm your earlier sentence.

THE COURT: All right.

At this time, I'll give Mr. Posey an opportunity to address the Court?

MS. YOUNG: Your Honor, may he come around to the podium?

THE COURT: Yes

MS. YOUNG: Thank you.

DEFENDANT POSEY: Thank you, Your Honor.

First, I'd like to thank you, Your Honor, and this courtroom, and your deputies, your clerk, and all your staff members for taking the time out to see me here today. I would like to thank the Prosecutor also for being here today for taking the time out of their busy schedules, and for agreeing to correct an error.

I'd like to thank my lawyer, Miss Young, for all

RECEIVED by MSC 4/12/2022 1:07:49 PM

the hard work she has put in for getting us here today. I truly appreciate everything you have done for me.

To my support team system there behind me, and the ones that couldn't make it here today, I love all yall. I truly appreciate yall being here and grateful for yall being here today. I know yall had to take off work or find babysitters for the kids and everything. I know yall was here for me and I truly appreciate that and I'm truly blessed to have yall in my corner.

Your Honor, I understand why we're here today for resentencing, and I know that there are guidelines that you still have to follow, but before you make your decision on the time that I would be leaving here with today, I just want you to take a few things into consideration.

I am not a career criminal. I am not a danger to society of any kind. I do have decision making -- poor decision making skills. Sometime I lack in thinking before I react, as my mother always tells me. I'm not putting the blame on my actions to anyone. I take full responsibility for all of the things I have done.

I will continue to try to grow. I'll continue to try to learn from this mistake I had created and I just hope that we can leave here today with something

RECEIVED by MSC 4/12/2022 1:07:49 PM

reasonable for both parties to feel is fair for both sides.

I know the things I have done is serious, and I know it caused harm mentally and physically to people. I can't explain how those people are feeling, but I know it's terrible. I know it's something I don't want to experience, and something I don't want my family to experience. I hate that the situation happened not even just because I'm going through this, but just because they've had to experience these things, my family had experience these things, and their family had to go through these things.

I want this court to know during the trial I did not take the stand. I did not make no statements, but my intentions was never to hurt anyone. I just wanted to say that things just got out of control. If I could change it, I would. I wish I could. If I would change it, I would.

But, I just wanted to let you know I lay at the mercy of this court, and I hope that I could leave here with something decent that says I'll be able to hug my mother again as a free man, be with my family, and live a normal life. Thank you.

THE COURT: All right. Thank you.

All right. The Court has had an opportunity to

RECEIVED by MSC 4/12/2022 1:07:49 PM

review the new Pre-sentence Investigation Report. I have heard from both Counsel, reviewed the sentencing memorandum submitted on behalf of Mr. Posey, and I've heard from Mr. Posey himself. And, in taking all of that into consideration, the Court will sentence Mr. Posey as follows:

While Mr. Posey may have had some improvement while he was in prison in his attitudes and his outlook on life, it doesn't change the fact that these offenses were committed while he was on parole after having previously served time for a similar offense, and the Court is not convinced that his sentence should be any different than what was imposed before. The prior sentences were well within the guidelines.

While, now they are advisory, you will be sentenced on Charges 1 and 3, assault with intent to commit murder as a third habitual offender to a minimum of twenty-two, to a maximum of forty years in the Michigan Department of Corrections. That's on Charges 1 and 3.

On Charge 5, the carrying a concealed weapon as a third habitual offender, a minimum of four, to a maximum of ten years.

On Charge Number 6, carrying a dangerous weapon with unlawful intent as a third habitual offender, four

RECEIVED by MSC 4/12/2022 1:07:49 PM

years to ten years in the Michigan Department of Corrections.

On possession of a firearm by a felon, a minimum of four, to a maximum of ten years in the Michigan Department of Corrections as a third habitual offender.

And, let's see. Then on Charge 9, which is the felony firearm second offense, the mandatory term of a flat five years.

On Charge 11, felony firearm, second offense notice, a mandatory flat sentence of five years.

On Charge 14, felony firearm second offense, a mandatory flat years of five years.

On Charge 15, felony firearm second offense, a mandatory flat five years.

Those sentences will be served first except with respect to the carrying a concealed weapon. That is not a predicate felony of felony firearm, so the felony -- I'm sorry -- so carrying a concealed weapon will start with the felony firearm sentences, and on those he would be entitled to a jail credit of 436 days.

All right. While the guidelines are merely advisory now, the sentences are still well within those guidelines.

And, all of these sentences will have to be -- whatever time he receives on his parole violation will

RECEIVED by MSC 4/12/2022 1:07:49 PM

have to be served first and then the sentences in this case will follow.

MR. ADAMS: Excuse me, Your Honor.

THE COURT: Yes?

MR. ADAMS: I'm not sure how you would want to indicate that on the judgment of sentence. I would recommend that it not be indicated on the judgment of sentence.

In McDaniel, 480 Michigan, 1162, the Michigan Supreme Court said that while it's true normally, a sentence is consecutive to something that a person is on parole, when there's a resentence they do get credit back to initial date even though there's been, in McDaniel and in our case I believe, there is no determination as to when he completed the first sentence. So, I don't think that there's any consecutive due to parole annotations that need to be indicated.

THE COURT: Well, that would seem to conflict with the recommendation of the Department of Corrections.

MR. ADAMS: Yes, it does. You're right.

In their Pre, they do say zero credit. I'd be glad to bring my copy of the McDaniel case to you.

THE COURT: What's the cite on that?

RECEIVED by MSC 4/12/2022 1:07:49 PM

MR. ADAMS: The cite on that is 480 Mich, 1162. It is designated as an opinion, not merely an order. It's a 2008 case.

And the facts in that were simply the Defendant was resentenced for a charge when he was on parole, and the original court in Oakland County gave him no credit, and the Court of Appeals affirmed that, and the Michigan Supreme Court unanimously said no, he's entitled to credit for time served on the void sentence while his appeal was pending.

THE COURT: He's received that credit?

MR. ADAMS: Yes. He has received that credit.

I was just concerned because you indicated that there was -- that this would be consecutive to the parole matter. That was all I was suggesting is that should not deprive him of his sentence credit.

THE COURT: No. He will be entitled to sentence credit.

MR. ADAMS: All right then.

THE COURT: I'm sure the Department of Corrections will advise us accordingly.

MR. ADAMS: Thank you, Your Honor.

THE COURT: All right.

We now have nine felonies instead of ten. And, nine times sixty-eight -- the State costs will be

RECEIVED by MSC 4/12/2022 1:07:49 PM

$612.00 for the nine felonies, $130.00 for the crime victims assessment, $1,300.00 for court costs. That figure has been arrived at as an average cost of processing cases through the Wayne County Circuit Court by the State Court Administrator's office.

Then I have to ask you this question, Mr. Posey, pursuant to the Court of Appeals. Would you be able to reimburse the Court in the amount of $400.00 for attorney fees?

DEFENDANT POSEY: No, sir.

THE COURT: All right. I won't impose that.

Appeal Rights?

MS. YOUNG: I have them with me, Your Honor.

For the record, I'm handing Mr. Posey a Notice of Right to Appellate Review and the request for appointment of attorney. I have two copies with me, one for him to sign receipt of notice, and the other for him to keep.

THE COURT: All right.

He still has a right of appeal in terms of these proceedings, correct?

MS. YOUNG: I've handed them to the deputy.

THE COURT: You have a right of appeal under these circumstances.

If you want to exercise that right, you may do so

RECEIVED by MSC 4/12/2022 1:07:49 PM

by signing and mailing back to this court, Mr. Posey, a copy of the form you're taking with you within forty-two days from today.

If you cannot afford an attorney to file a claim of appeal and brief of appeal on your behalf, the Court will appoint one for you and furnish that attorney with copies of all records transcripts necessary to make that filing.

Do you understand your appeal rights, Mr. Posey?

DEFENDANT POSEY: (No response.)

THE COURT: You have to answer out loud, sir?

DEFENDANT POSEY: Yes. I understand that.

THE COURT: All right.

Is there anything further at this time then for the Court's attention on the matter?

MR. ADAMS: Nothing, Your Honor.

MS. YOUNG: Yes, Your Honor, one more thing.

THE COURT: Yes?

MS. YOUNG: I just request on behalf of Mr. Posey if this court could delay the order to remit prisoner funds for one month, then that would give him time to get settled in the community adequately with family because when that starts taking out, it's a substantial hit on his own account. So, just to get -- so he knows where he is.

RECEIVED by MSC 4/12/2022 1:07:49 PM

In the MDOC, he may change locations within this next month as he gets settled. If this court could just delay sending that documentation for one month, that's the easiest way to do it because even if you send it with the notation, they don't start this for a month, but they start it the second they get it. So, I'm just asking for a one month delay in sending the order to remit prisoner funds to the Michigan Department of Corrections.

THE COURT: Pursuant to the policies enforced by the chief judge of this circuit, I cannot do that. A motion would have to be filed with the chief judge to have some relief from the imposition of costs and the withdrawal of monies from prisoner funds. You can bring that motion to the chief judge of the circuit court.

MS. YOUNG: Thank you, Your Honor.

THE COURT: All right. Thank you.

(At about 11:40 A.M. - proceedings concluded.)

- -

RECEIVED by MSC 4/12/2022 1:07:49 PM

CERTIFICATION OF COURT REPORTER

STATE OF MICHIGAN)
) ss.
COUNTY OF WAYNE)

I, REGENIA S. VEASY, CSR-2350, an Official Court Reporter in and for the Third Judicial Circuit Court, Criminal Division, State of Michigan, do hereby certify that the foregoing pages 1 through 38, inclusive, were prepared by me in the matter of DAMETRIUS BENJAMIN POSEY, Case Number 18-00074-01, on Thursday, November 7, 2019, and was reduced to typewritten form by means of Computer-assisted Transcription, and comprise a full, true and accurate transcript of the proceedings had in the above-entitled cause.

Regenia S. Veasy
Official Court Reporter

Feb. 13, 2020
Date