The Intuitive-Override Model: Nudging Judges Toward Pretrial Risk Assessment Instruments

Draft Prepared by:
Matthew DeMichele
Senior Research Sociologist
Center for Justice, Safety, and Resilience
3040 E. Cornwallis Road
Research Triangle Park, NC 27709-2194
919-541-6452
mdemichele@rti.org

Megan Comfort
Senior Research Sociologist
Behavioral and Urban Health Program
San Francisco, California
415-848-1375
mcomfort@rti.org

Shilpi Misra
Public Health Analyst
Center for Justice, Safety, and Resilience
RTI International, Hill 401
Phone: 919-485-2771
Email: smisra@rti.org

Kelle Barrick
Research Sociologist
Center for Justice, Safety, and Resilience
3040 E. Cornwallis Road
Research Triangle Park, NC 27709-2194
919-541-6854
kbarrick@rti.org

Peter Baumgartner
Research Data Scientist
3040 E. Cornwallis Road
Research Triangle Park, NC 27709-2194
919-541-5807
pbaumgartner@rti.org

Abstract

The development, implementation, and use of risk assessments are some of the most important and controversial issues facing criminal justice systems today. The recent push for judges to use risk assessments in their pretrial release decision making has met with resistance on several fronts, including from judges themselves. We conducted in-depth interviews with judges in a diverse set of courts that were using the Public Safety Assessment (PSA), a risk assessment developed by the Laura and John Arnold Foundation. Interview questions aimed to gain insight into how judges define and assess risk, as well as how they perceive the potential for bias and disparate impacts for communities of color in the use of risk assessments. In this manuscript, we frame our analysis with Guthrie and colleagues' *intuitive-override model* and suggest that risk assessment instruments can help judges engage both intuitive and deliberative models of decision making. Our findings indicate that judges perceive a tension when reconciling the actuarial aspect of the PSA and their experience-based inclination to learn about defendants' lives as a way of assessing risk through determinations of culpability and blameworthiness. We conclude by discussing the PSA through a review of legal scholarship that challenges risk assessments on ethical and moral grounds, and recommend the creation of researcher-judge feedback loops along with increased transparency of model development as key features to improve the accuracy, adoption, and understanding of risk assessments.

How do judges make pretrial release decisions? What influences judicial beliefs about risk and dangerousness? Judges regularly make difficult decisions about which individuals to release and which ones to detain during pretrial hearings. But, how do they make these decisions? What information do judges use? Judges, essentially, are making decisions about the probability of uncertain events – i.e., the likelihood of an individual making it to court and staying out of legal trouble. They review case files, assess criminal histories, and learn as much as possible about an individual prior to making their decision. Pretrial release decisions are usually made quickly and with limited information as judges make dozens of such decisions during a single court session.

To make these rapid decisions, judges are (often subconsciously) performing a series of intuitive calculations to predict the probabilities of how an individual will behave in the community. As an example, consider the simplification in which an overly lenient judge releases everyone, or, alternatively, an overly punitive judge detains everyone (by making release conditions onerous). Releasing everyone eases the burden on jails, alleviates individual problems associated with incarceration, and extends liberty. However, an unknown proportion of these individuals will fail to appear in court, some will commit a new crime, and a subset will commit a new violent crime. Alternatively, detaining everyone creates problems related to jail costs and crowding, ethical issues, and exacerbates collateral consequences related to incarceration.

These decision-making examples demonstrate key concerns when predicting pretrial outcomes. Releasing everyone results in false negatives as everyone is predicted to perform well, whereas detaining everyone increases false positives[1] as most individuals return to court and do not get arrested during the pretrial phase. All predictions, decisions, or assessments must balance these errors, but making so many rapid decisions with limited information requires a balance between

---

[1] The false positives, of course, cannot be measured in the hypothetical situation of detaining everyone because they are detained and as such do not have the opportunity to fail.

subjective/intuitive and objective/deliberative decisions. No doubt these decisions are rooted in judicial expertise gained from years of education and experience on the bench. Despite judges making qualified decisions rooted in their experience and knowledge of each case, there is growing evidence that "even highly qualified judges inevitably rely on cognitive decision-making processes that can produce systematic errors in judgement" (Guthrie, Rachlinski, and Wistrich, 2007: 3).

Recently, there has been a push toward more structure to pretrial decision making by using risk assessment instruments, but this push comes with some controversy. A recent ProPublica article challenged the use of risk assessments to make pretrial decisions by comparing error rates between predictions and outcomes among black and white individuals using the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) risk assessment instrument. They claimed that the use of the COMPAS equated to "machine bias" that resulted in "significant racial disparities" (Angwin, Larson, Mattu, and Kirchner, 2016). Flores, Bechtel, and Lowenkamp (2016: 45) responded to the ProPublica article by analyzing a similar dataset using a different statistical method, and came to a nearly opposite conclusion that there was "no evidence of racial bias."[2] These studies have set off something of a firestorm within the criminal justice research and practitioner communities as well as numerous media reports challenging the efficacy of risk assessments on ethical grounds.

The development, implementation, and use of risk assessments are some of the most important issues facing criminal justice systems. There are important concerns related to disparate impacts based on sex, age, and race, and experts are having trouble agreeing on the empirical measurement of bias (see Hannah-Moffat, 2013; Starr, 2014). One thing often forgotten in these

---

[2] Statisticians generally use four types of assessments to test for bias in an algorithm – i.e., error rate balance, calibration, predictive parity, and statistical parity. Chouldechova (2017) provided a third analysis of the COMPAS data and showed that differences in failure base rates by race make it impossible for these data to satisfy all fairness assessments. The ProPublica analysis assessed error rate balance (i.e., equal false positive and false negative rates across races) and Flores et al. (2016) assessed calibration (i.e., does a score of x mean the same thing for white and black defendants).

debates is the voice of the individuals that make decisions with the risk instrument.[3] The Laura and John Arnold Foundation (LJAF) developed the Public Safety Assessment (PSA) to assist judges to make pretrial decisions that has been adopted by more than 30 jurisdictions (including two state systems). We conducted interviews with judges using the PSA in a diverse set of courts to gain insight into how they define risk, assess risk, and perceive bias and disparate impacts for communities of color. The findings provide a crucial glimpse into how judges are thinking about pretrial risk assessment instruments.

The paper is arranged to first provide a thumbnail sketch of the use of risk assessment instruments within criminal justice systems. Next, we provide a brief discussion of the science of decision making to demonstrate the potential for systematic errors, especially when making decisions quickly with limited information. We frame the current study with what is referred to as the *intuitive-override model* and suggest that risk assessment instruments can help judges engage both intuitive and deliberative models of decision making (Guthrie et al., 2007). This suggests that most decisions judges make – especially pretrial decisions – are intuitive, fast, and rooted in their prior experience, but these more spontaneous forms of decision-making can be balanced or *overriden* with more objective criteria. Third, we describe the methods and procedures. Fourth, we present the findings by identifying emergent themes from our interviews in which judges stressed the tension they face when reconciling the actuarial aspect of the PSA as they try to learn about defendants' lives. The interviews showed that judges rely on criminal background to assess culpability and blameworthiness by reviewing criminal background and prior violence, and mixed views on the potential for bias against people of color.

---

[3] Miller and Mahoney (2013) used a national survey of community corrections staff and Viglione, Rudes, and Taxman (2015) used interviews and observational data to report that probation officers comply with requirements to complete risk assessments, but rarely used the assessment scores to make case management and supervision decisions. Our paper specifically focuses on interviews with judges.

We conclude the article by reviewing the PSA through legal scholarship that challenges risk assessments on ethical and moral grounds (Tonry, 2014). There is general agreement that risk assessments need to meet both empirical and ethical standards, and ascribed characteristics such as race and gender are left out of most prediction models even though they might improve predictive validity (Monahan, Skeem, and Lowenkamp, 2017; Tonry, 2014). Legal scholars have assessed the merits of risk assessments to make sentencing and parole decisions, with little treatment of their use during pretrial. In the end, we recommend the creation of researcher-judge feedback loops, and the need to increase in the transparency of model development as key features to improve the accuracy, adoption, and understanding of risk assessments.

Criminal Justice and Risk Assessment

Criminal justice and legal professionals assess risk on a regular basis. A police officer assesses risk when deciding to administer a citation instead of arresting someone. Parole board members assess risk when deciding to release someone. Judges, of course, assess risk when deciding whether someone can be released pretrial – when they are still considered innocent – or if they should remain in jail while awaiting trial. These professionals make such decisions many times throughout a given day. There are nearly 12 million jail admissions annually (Minton, 2015) and criminal justice professionals need to make quick decisions and these decisions have important ramifications for each person's life. Risk assessment instruments are intended to make these decisions more systematic and accurate, and easier for practitioners.

Risk assessments are based on clinical judgement or actuarial practice (Gottfredson and Moriarty, 2001). Clinical judgements are often referred to as first-generation assessments that are based on intuition with assessment of risk based solely on subjectivity or "gut feelings" derived from education and experience (Bonta, 1996). The purely actuarial approach, or second-generation of risk

assessments rely on a more formal, statistical model of risk that should provide more consistency and uniformity in risk classification (Barbaree et al., 2006; Harris and Rice, 2003). It is common for many criminal justice actuarial risk assessments to allow professionals to adjust scores based on nuanced information about the individual.

The use of risk assessment instruments in criminal justice settings is not new. Several sociologists assisted parole boards and prisons develop predictive instruments starting in the 1920s. Burgess (1928) worked with the Illinois State Parole Board to develop a parole release instrument that relied on an additive binary assessment instrument of 21 factors to predict which offenders were most likely to succeed and fail on parole. Numerous studies and meta-analyses have found that decisions guided by statistically-derived tools provided a more accurate result than clinical assessment (Groves and Meehl, 1996; Groves, Zald, Lebow, Snitz, and Nelson, 2000). Groves and Meehl (1996, p. 293) stated that the "conclusion was clear that even a crude actuarial method…was superior to clinical judgment in accuracy of prediction."

Risk assessment instruments are developed for specific jurisdictions and specific phases of the criminal justice system. As a warning, Skeem and Lowenkamp (2016) made clear that jurisdictions will face potentially large error rates and inconsistency when using assessments for a different phase than they were intended. The use of pretrial risk assessments has been in effect since the early 1960s, claiming to objectively assess the public safety and failure to appear (FTA) risks of releasing defendants from jail. The Pretrial Justice Institute (2015) estimated that there are about 10 percent of pretrial agencies using pretrial risk assessment instruments. This means that most pretrial release decisions are made without the guide of an actuarial instrument.

In this paper, we seek to contribute to understanding the judicial decision-making processes during the pretrial phase by examining judges' perspectives on and use of risk assessments. In the following, we lay the groundwork for how such assessments might fit into the framework of judicial

decision making overall by discussing the intuitive and deliberative decision-making frameworks. Judges complete education and training in which they gain a fundamental knowledge about the law, procedural rules, and legal processes. Pretrial researchers have pointed to the difficulties involved with making release decisions due to the speed and volume in which these decisions are made (Sacks, Sainato, and Ackerman, 2015; Demuth, 2003). Research on human judgement and choice demonstrates that judges – similar to engineers, accountants, military leaders, and others – rely on several cognitive shortcuts to process information quickly when making decisions under uncertainty (Guthrie et al., 2001).

Judicial Decision Making

Nearly 90 years ago, the legal scholar, Jerome Frank (1930), observed that judges base their decisions on hunches, and "whatever produces judges' hunches makes the law." Frank recognized the importance of judicial subjectivity and intuition when making decisions. Much has been written about judicial decision-making, and the tensions between intuitive and deliberative decision-making. Legal scholars, for the most part, fit into one of two camps with legal formalists suggesting that judges apply the legal rules in a logical, mechanical, and deliberative manner whereas legal realists say that judges make decisions through intuition and only later rationalize with deliberative reasoning. We suggest that neither of these perspectives are entirely correct, with cognitive science showing that both operations are at work (Guthrie et al., 2007).

There is a large body of psychological and behavioral economic research showing that human decisions are made with dual processing mechanisms (e.g., Tversky and Kahneman, 1974; Kahneman and Tversky, 1979; Kahneman, 2011). Dual process models of cognition divide cognitive processes into two systems to differentiate between intuition and reasoning. There are several versions of dual process models, but each version distinguishes between the cognitive processes that

are "quick and associative from others that are slow and rule-governed" (Kahneman and Frederick, 2002: 51). Stanovich and West (2002) labeled these System 1 and System 2 to differentiate cognitive operations by their speed, control, and information. To put it simply, dual process models suggest that human beings make decisions using automatic, intuitive, and non-reflective processes (i.e., System 1), and deliberate, thoughtful, and rational processes (i.e., System 2). These systems do not operate in isolation of one another, but rather new information is processed, stored, and remembered through System 2 learning processes. With complex information moving from System 2 to System 1 as individuals acquire a degree of proficiency and skill – essentially, experts rely on System 1 automatic processing as they gain pattern awareness (Kahneman and Frederick, 2002). There are numerous examples of how novel information becomes engrained and hard wired, as we rarely need to engage in much effort when driving a car, reading a book, or walking.

These cognitive processes are at work when judges are making decisions. Judges learn through experience on the bench as they interact with defendants and court staff as well deliberative study of the law. A brand-new judge,[4] for instance, will engage in deliberative and effortful cognitive processes to learn, remember, and apply their knowledge of legal rules and courtroom cultural norms. Over time, a seasoned judge will have engrained this knowledge of the law and normative behavior to allow for most judicial decision making to move from System 2 to System 1. But, as we rely more on intuition for decision making, we run the risk of making errors because these decisions are made quickly, with little reflection (Frederick, 2002). There is a large body of psychological research on heuristics and biases (Kahneman and Frederick, 2002) that show that people base decisions using mental shortcuts, cues, and stereotypes. This applies to judges because they could erroneously, and unwittingly, introduce bias through acquired stereotypes in which they come to

---

[4] This description of a judge is only meant as an example because judges very often will have been lawyers beforehand, attended law school, and have some general idea of how the law operates.

automatically associate being black, male, or young with criminality and violence (Hoschild and Weaver, 2007). Papillon (2013) provided a thorough review showing the connection between cognitive science and legal decision making to demonstrate the different neurological processes at work that associate (due to specific brain functions) being black with fear, threat, and aversion.

This line of inquiry has been applied to the legal field to study judicial decision making. Guthrie et al. (2007) adapted Kahneman's dual process model to develop an intuitive-override model for judges. Through a series of cognitive experiments with trial judges, Guthrie et al., (2001, 2007) found that judges rely on similar cognitive heuristics (e.g., anchoring, statistical errors) that result in common decision errors (e.g., reliance on arbitrary references, ignoring trends). Additional experiments tested judges' reliance on intuitive versus deliberative decision making and found that judges are as reliant on intuition as other populations (e.g., physicians, engineers). The intuitive override model starts from the assumption that judges (as all humans) generally engage in intuitive decision making, but their intuition can be disrupted to create an opportunity for judges to reflect on decisions as needed.

The essential argument of the intuitive override model is that "judges should use deliberation to check their intuition" (Guthrie et al., 2007: 5). This approach fits with other in the legal scholars identifying techniques to combat implicit bias in the courtroom (Kang et al., 2012). We do not test Guthrie et al.'s intuitive-override model. Instead, we use this model as a framework to view and understand a series of judicial interviews about the use of a risk assessment instrument to make pretrial decisions and to offer recommendations more broadly about judicial use of pretrial risk assessments. The judiciary are equally susceptible to common psychological heuristics that can produce systematic errors in judgement that result in bias and disparate treatment. The PSA and risk assessments more broadly are a potential tool that judges can use to question their hunches.

Methods

  The analyses are based on interviews conducted with judges in three geographically distinct jurisdictions in which judges were using the PSA. The interviews are part of a larger project to validate the risk assessment instrument and understand the implementation and use of the risk assessment instrument. The purpose of the interviews was to better understand judges' perceptions of the usefulness of the risk assessment instrument to make pretrial decisions.

Public Safety Assessment (PSA) Design and Use

  Before detailing the interview procedures, we briefly describe how the PSA was developed[5] with direction and funding from LJAF and how jurisdictions use the PSA. The PSA was developed using nine datasets from seven states (i.e., Colorado, Connecticut, Florida, Kentucky, Ohio, Maine, and Virginia) and two datasets from the Federal Court System to calculate probabilities of failure to appear in court (FTA), new criminal activity (i.e., any new arrest), and new violent criminal activity (i.e., these definitions are developed to fit each specific jurisdiction).[6] Jurisdictions implementing the PSA received technical assistance and training to describe the research used to develop the instrument (provided by Luminosity or Justice System Partners), provide detailed instructions for completing the PSA, and offer ongoing support during implementation. The implementation team focused on providing jurisdictionally tailored training and technical assistance to ensure that the instrument could be successfully implemented in each jurisdiction. The implementation team, for

---

[5] The authors of the current paper were not involved in the development and validation research used to develop the risk assessment instrument. We are conducting a broader research and validation project of the risk assessment instrument in which we are collecting available datasets used for development and validation by the risk assessment instrument development team. The current analyses do not assess the validity of the risk assessment tool or the procedures used to develop the instrument. Instead, we seek to understand judicial views about the use of the instrument.

[6] The instrument development team processed these datasets to identify the predictors of each of the three outcome variables. They used a series of statistical techniques (e.g., logistic regression, contingency tables) that produced hundreds of effect sizes. The effect sizes were averaged, and were restricted to variables that were at least one standard deviation above the mean effect size. Further analyses were conducted to identify the best effect sizes and operationalization in which each predictor variable had at least a 5 percent increase in likelihood of failure to appear or new criminal activity. The new violence criminal activity flag used a variable selection criteria of doubling the probability of failure when the item was included in a model (this paragraph is adapted from unpublished materials by Luminosity).

example, would learn specific information about each jurisdiction's capacity to collect the needed defendant information, identify appropriate communication procedures to share the results of the risk assessment with judges, defense attorneys, and prosecutors.

Pretrial officers complete the PSA prior to first appearance. Pretrial officers identify eligible defendants[7] for the pretrial release instrument using administrative data and conduct a thorough review of criminal history records. The PSA includes eight criminal history/conduct factors and a categorical age factor. Below are the three outcomes and each of the factors:

- Failure to appear: pending charge at time of arrest, prior conviction, prior failure to appear within two years, and prior failure to appear longer than two years.
- New criminal activity: pending charge at time of arrest, prior misdemeanor conviction, prior felony conviction, prior violent conviction, prior failure to appear within two years, prior sentence to incarceration, young age at current arrest.
- New violent criminal activity: pending charge at the time of arrest, prior conviction, prior violent conviction, current offense violent, and current offense violent * young age at current arrest.

The factors are weighted and converted to separate FTA and new criminal activity scales that range from 1 to 6, and a new violent criminal activity flag (i.e., binary indicator of yes/no). The new violent criminal activity flag is used to either recommend release or more restrictive conditions including detention for the defendant.

The FTA and new criminal activity scale scores are converted into recommendations for each defendant that a judge may choose to follow (or not). The decision-making framework provides jurisdictionally based guidance on the recommended nature of release for an individual, which can range from release on own recognizance, various levels of supervision, and recommend detention. The decision-making framework is a key element of the risk assessment instrument to assist judicial decisions. The new violent criminal activity score produces a binary indicator as a

---

[7] The implementation of the risk assessment instrument allows local jurisdictions to identify case types and charges that are of such a nature that the jurisdiction excludes them from the assessment. Outside of these few crime types, all individuals booked in jail on a new crime are scored with the PSA.

violent "flag" to signal to judges that the defendant has a high potential for violence, and this case should be reviewed more carefully before making the release decision. The specific way the risk assessment instrument is completed varies to fit each jurisdiction's standard operating practices and courtroom culture.

Site Visits

Sites visits were arranged through initial email inquiries to explain the purpose, goals, and procedures for the visits to determine interest and availability. Once agencies indicated interest, the research team arranged a phone conversation to discuss logistical details. The authors divided the three sites to have one senior researcher act as the main liaison for each site. Each visit lasted about 2 days, and included interviews with judges, public defenders, prosecutors, and pretrial officers responsible for completing the risk assessment instrument. During our site visits, the research team observed pretrial officers accessing a series of databases needed to complete the PSA, and we observed pretrial hearings in each jurisdiction.

Three sites were selected for visits because they were "early adopters" of the PSA. These sites included an east coast city of about 1 million, a Midwestern city of about 500,000, and a west coast city of about 250,000. Although we do not suggest that our methods provide us with a nationally representative sample, we do, nonetheless, find the interviews to provide a unique opportunity to consider the complicated nature of assessing risk and judicial decision making. Interviews were audio recorded, and recordings were subsequently listened to multiple times by multiple researchers, with extensive notes taken by an analyst. Notes were reviewed with the analyst and the interviewers, and in an iterative process, salient themes were identified through discussion, developed in writing, and then compared against the data and discussed further. The following sections describe the continuum for which judges deconstruct narrative, assess risk, and use criminogenic factors in the consideration of public safety.

Although our site visits included interviews with other court professionals (e.g., prosecutors, public defenders), we only report the findings from the judicial interviews. The purpose of these interviews was to assess the use of the instrument by judges -as it is a tool to inform judicial decision making – because predictive tools are ineffective if the intended users do not understand them or use them. An alternative approach to understanding judicial use of the instrument is to compare concurrence rates to measure how often judges implemented the decision suggested by the instrument. These data did not exist in a consistent fashion in our sites, and with this portion of our study, we wanted to understand something more nuanced. That is, we wanted to address the gap in research about judicial perspectives about risk, assessments, and related biases. These are crucial phenomenon to understand when it comes to judicial decision making.

The interviews were semi-structured to frame them around key themes, while allowing judges the freedom to expand as needed. The interviews were conducted by 1 to 2 interviewers and lasted between 45 minutes and 90 minutes. The main research questions guiding the interviews were:

- How do judges define risk?
- How do judges view the use of risk assessments to assess risk?
- What are judicial interpretations of racial and ethnic bias in risk assessments?

Findings

In this paper, we explore an inherent tension that emerged in the analysis between the "subjectivity" of judicial discretion (a story-based assessment) and the "objectivity" of the risk assessment instrument (a numbers-based assessment). Below we discuss our findings about the information that judges seek when considering a case, and how they interpret the information provided by a risk assessment tool as fitting within the broader landscape of judicial decision making. We then present judges' perspectives on pretrial risk assessments and racial/ethnic disparities, highlighting the diversity of their opinions and approaches. We conclude with a

discussion of the implications of this tension for judicial override of risk assessment recommendations and consider factors that judges highlight as important when being informed about, trained on, and supported in the use of a risk assessment.

Pretrial risk assessments within the context of judicial discretion

Judges develop their own roster of information they seek when considering a case, and there is variability in the factors they value. When asked in interviews what elements were salient for them, judges described wanting information about a defendant's criminal background and previous violent offenses, whether the current charges involved weapons or physical injuries, and statements from the victim. With public safety foremost in their minds, they were on the lookout for aggravating factors like repeated failures to appear (FTA), increases on the violence scale, charges of attempted murder or assault with a deadly weapon, and whether someone was on felony probation and committed a new offense.

Judges also referred to factors based on their professional experience, including what they had seen in other cases, mistakes made, and lessons learned. As part of drawing on their experience, some judges referenced the importance of local culture, not only in understanding the issues that affect the population but also understanding the cultural values associated with a jurisdiction.

When weighing information and making decisions, judges indicated that they were interested in the ways that different factors compounded, corroborated, or cancelled out each other. For example, did someone with a high number of FTAs also have a history of substance addiction or mental health issues? Was someone repeatedly showing up in court being charged with the same crime? Looking at the interactions of multiple factors to get a sense of an underlying story was important, especially when judges were determining whether to refer the defendant to a diversion

program or pretrial release under conditions of supervision such as electronic monitoring or routine drug testing.

Unsurprisingly, judges' perceptions of the utility of the risk assessment instrument were strongly shaped by the overlap between the information they considered valuable when exercising their decision-making and the information used to complete the risk assessment. Within the construction of juridical stories, all judges interviewed found value in the risk assessment instrument to some degree. Those who were most favorable viewed the risk assessment instrument as an expeditious means to synthesize the information that they felt was important, namely FTAs, charges involving violence, and recent convictions. For these judges, there was appeal in a tool that was "not based on subjectivity and sympathy" and that allowed them to quickly assess "on a busy calendar… [how] to zero-in on what can be done for each specific person."

Notably, judges who were highly favorable of the assessment instrument concurred not just with the information that it took into account, but the recommendations that it provided. In this regard, these judges expressed that they thought the assessment instrument was a practical and useful tool for others, especially their younger and less experienced colleagues, but that in their own use of it they tended to see it more as confirming the decision they had arrived at through their judicial discretion. As one judge stated,

> "For most judges who don't feel confident to go deeper, that [the assessment instrument] is fantastic. If we are talking about one size fits everybody, wow, this is great, right?"

While speaking enthusiastically about the assessment instrument and expressing confidence in its reliability, this judge also indicated that he did not perceive a need to rely on the tool in his decision-making. In addition, he noted instances in which he would override the risk assessment recommendation, such as when the defendant had known mental health issues or there was a

weapons charge, stating that the judiciary's role was to "connect those dots that a pretrial services report does not understand."

Judges who were more reticent about the risk assessment instrument tended to frame their perceptions of the tool in terms of the information it did and did not incorporate. One judge spoke broadly about the concept of risk, and in doing so listed three factors that are not addressed by the assessment instrument:

> "Risk has changed over time because personal experience and training has caused people to think about predicting risk – substance abuse, housing stability, connections to the community – and making recommendations for conditions of pretrial release."

Similarly, in jurisdictions where a different pretrial risk assessment instrument had been used previously, legal actors had become accustomed to associating certain factors with risk, and the removal of those factors felt disorienting and counterintuitive to judges, as if refuting years of prior practice. In these discussions, judges acknowledged the tension between the research validation of the new risk assessment instrument and their experience on the bench. As one judge commented:

> "It's hard to wrap your head around releasing someone with a felony II burglary, but all of the factors and data across a million cases say that this person isn't coming back with another burglary and there isn't a risk to public safety or court."

When considering the research, some judges raised questions as to whether the factors that were found to be predictive in one jurisdiction were accurate in predicting risk in other jurisdictions, again valuing their own expertise and local knowledge. As one judge remarked, "it is one thing to say that the research says that these things are predictive, and here's the tool [but] we need the validation process to use local data."

In addition to concerns about the types of information used to generate the risk assessment instrument, which led to questions about what elements were missing, some judges also expressed

skepticism about how the absence of information might obfuscate the underlying story that they were looking for in their decision-making process. One example that was brought up multiple times in interviews were the cases of defendants who cycle frequently through courts and have a high probability of FTA. One judge explained that public nuisance offenders often fit this category and were scored as very high risk and not recommended for release, which struck her as counterproductive. She was inclined to override these recommendations and stated that judges needed alternative resolutions at pretrial release.

Several other judges raised the example of domestic violence cases, saying that when these cases came before them they wanted more information than the risk assessment instrument provided, specifically violent criminal history, domestic violence history, police reports, victim's statements, and lethality assessments. In cases when the victim indicated being fearful of the defendant or there was complex information in the probable cause statement, judges consider overriding the risk assessment instrument's recommendation based on this information.

Among all the judges interviewed, there was a tendency to characterize the risk assessment instrument as "one tool among many." Those who were highly favorable of the PSA were still inclined to consider recommendations in the context of their own judicial intuition and experience, and would request information that was not included in the risk assessment instrument when they deemed this to be necessary. As one judge remarked, "It's important to understand that it's just a tool and that judges are the definitive answer." In his view, more work was needed to enforce that the evidence-based tool was intended to support judicial decision-making, and not to replace that process.

Racial and ethnic disparities: Differing perspectives

Of the judges interviewed, there was general agreement that communities of color are disproportionately represented in the criminal justice system, and many judges acknowledged the potential for bias in judicial decision-making. One judge who identified himself as coming from a heavily policed community expressed that it was important to have judges from minority racial and ethnic groups and from low-income backgrounds to counteract bias:

> "How do we view people? Do you see kids walking down the street and go for the automatic door lock?"

This judge also thought that a more representational judiciary would help to build the credibility of the justice system, noting that it was difficult to ask people to "have faith" in the courts when the judges "look nothing like them."

Judges overwhelmingly supported risk assessment tools for providing a means of making release decisions in a way that was separated from knowledge of defendants' physical characteristics, seeing this as "a benefit to a tool like this because people are not always aware of their prejudices." Indeed, when asked about how the risk assessment instrument might help reduce disparities, the majority of judges stated that because the tool takes only objective factors into consideration, it would either minimize disparities or have no influence on them. Among people holding these views, the risk assessment instrument was seen as useful in reducing bias in the criminal justice system. One judge suggested that bias is not an in issue in his jurisdiction nor does the risk assessment instrument create any disparities. He stated:

> "I do not think that disparities are a concern in pretrial release. The racial or sexual element is not an issue because it's an objective system. I've never given it any thought. In [jurisdiction name], I don't think it's an issue. I don't think the PSA makes any kind of distinction that would lend itself to creating disparities."

Another judge emphasized that when considering the potential biases from an instrument, it is important to realize that prior judicial decision-making criteria might have generated disadvantage for communities of color and the poor. She suggested that prior decision-making included factors related to socioeconomics in which the poor and homeless were less likely to be released:

> "The tool (PSA) does not have a slant or bias in the recommendations... [Before the PSA] We used to assess ties to the community and where people work but we don't do that anymore. They have 2,500 homeless people in [jurisdiction name]."

The essence of this judge's comments was to point out that in this jurisdiction they are moving away from using criteria such as homelessness to keep people in jail pretrial. The adoption of the risk assessment instrument, although not the only reason, has contributed to less attention to variables such as housing stability and employment status directly influencing release decisions.

The judges varied on their perspectives about racial/ethnic bias within their jurisdiction and related to the risk assessment instrument. One judge disagreed with this perspective, arguing that even "objective" factors contain a subjective context and that people's criminal records are the result of socioeconomic, racial/ethnic, and gender dynamics that affect the likelihood of an individual being arrested, prosecuted, convicted, and detained. In her view, quantifying risk factors such as "prior conviction" is of concern, given that communities of color are frequently heavily policed and "men of color are more likely to be arrested, stopped, searched, and charged for even minor drug-related offenses, [and therefore] they are more likely to have that risk factor on a subsequent offense." Another judge suggested that their decisions need to consider more nuanced perspectives because:

> "It's important to humanize defendants and that should be the role of the judicial team. It's important to educate judicial officials about the prevalence of low-risk people in jails and in pretrial. This requires education around disproportionality and how certain practices target people of color."

These judges used the logic described above of trying to construct a story beneath the risk factors, noting that a judicial decision could take into account that white people and African Americans have been found to use drugs and commit other infractions at similar rates, and therefore an African American person's conviction history might reflect more about bias than about criminal behavior. Likewise, this judge noted that the intersection of race/ethnicity and low economic status could produce scores that reflected lack of access to resources (e.g., transportation to the courthouse, resulting in an FTA) rather than risk to public safety.

Obviously, the judicial interviews demonstrate variation in perspectives about bias in the criminal justice system and whether risk assessment instruments contribute to any bias. Regardless of this variation, we did learn that these judges, for the most part, recognize the need to stem potential bias and they want to learn more about how risk prediction can contribute to their decision-making.

How do Judges Understand Prediction? They need/want more information

The limited understanding among judges and risk assessment tools is not due to a lack of institutional trust. Rather, there is a need for continuous education and learning to understand what the risk assessment instrument can provide, including a realistic contextualization of the likelihood of errors. One judge framed the nature of the likelihood for errors in both judicial intuition and the risk assessment by stating that "judges are right 50% of the time and the tool is right like 60% of the time, and the tool and judicial discretion is right about 80% of the time." Although we did not assess the changes in prediction errors when using the instrument, another judge described her desire to

understand the nature of the research by stating that it would be helpful to see probabilities of positive outcomes (emphasis added):

> "It would not be helpful for judges to see actual probabilities of people for an FTA or a new crime. People with these characteristics have a certain percentage for a FTA. I would rather see people with this X risk score *succeed* in appearing in court a certain percentage of the time, just to reinforce the predictive relationship."

Judges stressed the importance of presentations to new users and the substantive value in in learning more about the PSA. The judges expressed a need to develop a knowledge base about what goes into developing a risk assessment tools. Unique to jurisdictions, implementing data-driven research is the collaboration and coordination among probation, pretrial services, and judges through statewide meetings, to reinforce the importance of the tools, particularly for new judges cycling through first-appearance court. Meetings have contributed to judges feeling more equipped to receive guidance from the tool without the tool representing a threat to exercising judicial discretion. One judge described his experience:

> "…probation continues to [make] constant refinements [to the use of the PSA that] are helpful. It's important to have meetings and show stakeholders what's going on, there is a lot of buy-in. It's more difficult when people are saying that they don't know…when they do not go to meetings."

The point this judge was expressing is that judges (and other courtroom professionals) need to participate in trainings about pretrial risk assessments. Without these trainings, judges will be unaware of the potential for these tools, and they will not understand how to interpret and apply the recommendations from the decision-making framework.

The judiciary need a realistic understanding of the predictive ability of the PSA and to know exactly what is being predicted. It was suggested that intentionally targeting presiding judges in first appearance court is critical to keep individuals abreast of research, the tool itself, predictive factors,

validation from jurisdictions, the value of objective criteria versus subjective criteria, education on

disproportionality and implicit bias, and an institutional approach to the risk assessment instrument.

Moreover, judges expressed an interest to see how predictive the instrument is by using local data to

show the results from the instrument and the recommendations – e.g., how accurate is judicial

intuition and the instrument? Another judge summarized the way in which he perceives the use of

pretrial risk assessment instruments:

> "It's important to understand that it's just a tool and remember that judges are the definitive answer, and not to overly rely on a piece of paper. I'm glad that there is a tool that has evidence-based research."

In general, these interviews demonstrate that judges understand that risk assessment instruments can

support, clarify, and assist in judicial decision-making. But, they want more (continual) information

to demonstrate the accuracy or improvement in outcomes.

Discussion

The PSA was intended to offer judges standardized and objective information to inform

judicial decision making. The PSA was not to replace judicial decision making, but rather offer

recommendations using criterion related to pretrial outcomes. Pretrial risk assessments, on their

face, fit with Guthrie et al.'s intuitive override model to intervene in (System 1) automatic cognitive

processes. However, the PSA and risk assessments more generally need to meet certain moral and

ethical criteria. Risk assessments need to identify factors that are both highly predictive of the

outcomes and do not increase likelihood for disparate treatment. [8] Some judges articulated grappling

with the potential for bias in the criminal justice system, identified the potential for biased intuitive

---

[8] There is also a robust literature on the design of fair algorithms mostly outside of criminological and criminal justice fields. Data scientists and statisticians (e.g., Chouldocova, 2017; Corbett-Davies, Goel, Pierson, Huq, and Feller, 2017) have shown the inherent tension between maximizing predictive validity and equality across subgroups. There is "no algorithm that can simultaneously achieve calibration, error rate balance, and positive class balance (Corbett-Davies et al., 2017: 3). Instead, the quantitative literature demonstrates mathematical realities that there are tradeoffs that are "matters of values and law…not matters of science" (Berk et al., 2017: 34).

judgements, and viewed the PSA as a way to challenge those biases. The ProPublica and Flores et al. (2016) debate highlights the gap in understanding and consideration about bias and disparate treatment stemming from pretrial risk assessment.

Risk assessments have routinely been challenged on philosophical, ethical, and moral grounds to make sentencing and post-sentencing decisions. Criminologists were critical of the growing reliance on risk assessment instruments, with Feely and Simon (1992) arguing that risk assessments were managerial techniques that would inevitably increase the punishment of vulnerable populations. Garland's (2001) oft-cited term of mass incarceration signaled how correction practices (including risk assessments) were concentrated on specific groups – e.g., the poor, communities of color – and would invariably lead to increasing disparities. One of the harshest critics of risk assessments is the legal scholar, Michael Tonry (2014), who argued against most risk assessments out of concerns for respect, dignity, and fairness for the accused. Instead, of risk, Tonry argued that traditional jurisprudential factors of individual culpability, blameworthiness, and severity of the offense should be the primary concerns. Harcourt (2007) argued that risk factors are so highly correlated with race that risk scores are proxies for race, and others suggested that risk assessments merely provide a scientific veneer and discourse to support disparate treatment of people of color (Hannah-Moffat, 2013; Starr, 2014).

The bulk of the concerns about disparate impact stemming from risk assessments are related to ascribed statuses of age, gender, and race. The PSA includes a risk factor for young age at time of the arrest (for NCA), and whether the person was young and the current offense is violent (NVCA). Tonry (2014) identified age as a counterproductive risk factor because it seeks to harshly punish young people as they are continuing to develop socially and cognitively. Monahan, Skeem, and Lowenkamp (2017) found that including age in the federal Post-Conviction Risk Assessment (PCRA) overestimated the arrest rates for older individuals and underestimated arrest rates for

young individuals. Parsing the inclusion of age in the PSA is difficult because younger individuals have higher criminal propensity, but aging is known to promote desistance. Tonry (2014: 171, italics added) is unwavering in saying that "Ascribed characteristics for which *individuals bear no responsibility*, such as race, ethnicity, gender, and age, should not be included." Monahan et al. (2017: 200) point to some of the complexity around age as a risk factor because "Youth…both "diminishe[s] culpability" for past crimes (*Roper v. Simmons, 2005*) and enhances risk for future crime." There is a lack of clarity on setting standards about the inclusion of age on risk assessments, especially during the pretrial or sentencing phases.

Besides age, the PSA includes several criminal history factors, prior failure to appear, and whether the current charge is for a violent offense. These risk factors are commonly used in pretrial assessments and similar to input factors used to develop sentencing guidelines (e.g., Minnesota, Pennsylvania) related to prior criminal record and offense gravity scores (e.g., severity of the charged offense). Harcourt (2007) takes a strong stance against using criminal history as a risk factor because communities of color are over policed, charged, and sentenced relative to white neighborhoods. In one of the most comprehensive analyses of criminal history and sentencing, Frase, Roberts, Hester, and Mitchell (2015) described that some legal theorists suggest that prior criminal activity does not make one more culpable for the current offense (Hessick and Hessick, 2011), whereas more retributive theorists view repeat offending deserving of harsher punishments (Mahon, 2012). Frase et al., (2015) described the use of criminal history scores to make sentencing decisions, but there is a lack of guidance about the ethical implications of using prior convictions to make pretrial decisions.

The pretrial phase operates according to different legal rules than sentencing, in *United States v. Salerno*, the US Supreme Court upheld the use of dangerousness and flight risk as appropriate

factors to consider when making pretrial detention decisions.[9] In Gouldin's (2016) legal review, she

traced the historical trajectory of pretrial rules, bail reforms, and pretrial risk assessments to show

how dangerousness is poorly defined, and it is often conflated with flight risk. The PSA is unique

because it uses three algorithms to assess the likelihood of FTA, NCA, and NVCA, and combines

this information to create recommendations based on local practices. Unraveling the distinction

between dangerousness and failure to appear is complicated as someone with several failures to

appear because they lack transportation, have a drug addiction, or simply forgot about the court date

pose very different concerns than someone who fails to appear because they have left the

jurisdiction or are otherwise intentionally evading the court.

Conclusion

Everyday judges make thousands of pretrial release decisions about the nearly 12 million

individuals arrested each year. Pretrial is the beginning of the criminalizing process (e.g., Garfinkel,

1956), and, whether an individual is detained or released can have major implications for their life.

Being detained prior to trial can have negative effects on an individual's ability to maintain

employment, meet familial responsibilities, and may even contribute to negative health effects (see

Schonteich, 2011). Pretrial detention is intended to be used sparingly, but national pretrial detention

rates have grown from around 50 percent to 63 percent from 1990 to 2015, and an estimated 95

percent of the growth in jail populations since 2000 is due to the increased proportion of those

confined in jails being held pretrial (Minton and Zeng, 2016). Nationally, African-Americans are

detained at about three times their presence in the population, as about 35 percent of jail detainees

are black, whereas about 12 percent of the US population are black (Minton and Zeng, 2016).

---

[9] Although Salerno allowed for pretrial detention decisions to include concerns of dangerousness or flight, there are several variations across the states, with New York and New Jersey not allowing for detention based on dangerousness (see Baradaran and McIntyre, 2012).

The reasons for this disproportionality are beyond the scope of this paper, but most pretrial detention decisions are made without the aid of a risk assessment instrument. Instead, the bulk of these decisions are made relying on judicial intuition, legal knowledge, courtroom norms, and local cultural determinants. These decisions are aided by heuristics or mental schemas that ease cognitive operations involved in decision making.[10] We briefly sketched out how humans come to rely on automatic cognitive processes when making most decisions, but these automatic assessments can be wrong as they draw on potentially faulty associations, imprecise pattern recognition, and stereotypes. These cognitive errors have major implication for pretrial decision making especially regarding race, gender, and class, with judges potentially associating blackness, maleness, and poverty with aggression, aversion, and hostility (Papillon, 2013). The point is that if judges – as do the rest of us – have the potential to draw erroneous associations with various ascribed statuses can risk assessments override or provide a cognitive speed bump, if you will, to these automatic processes (Guthrie et al., 2007).

Journalists, scholars, and legal actors have questioned the potential unfairness in risk assessments. A lively debate has emerged recently about the fair use of risk assessments to inform decisions about pretrial release or detention. Similarly, psychological researchers have found that humans rely on intuition to make most decisions – even seemingly complex decisions such as pretrial decisions – that rely on error prone heuristics. Heretofore, however, we have heard little from the legal actors responsible for using the recommendations generated by the PSA. The current paper reveals at least three central findings from the judicial interviews. First, judges struggle to balance the importance of subjective (extra-legal factors) and objective criteria provided in the PSA – as they want to know more about the story behind the defendant (e.g., to humanize them).

---

[10] We do not assess the merit of pretrial risk assessments to reduce pretrial populations of disparities. Kleinberg et al. (2017) used a simulation method that found that a risk assessment could reduce jail populations without any increase in pretrial failures.

Second, judges want (and need) to know more about the PSA (or other risk assessments being used) to understand how it was developed, how the recommendations came about, and to learn more about accuracy and errors related to the PSA. There is a general call throughout the field to provide transparency about the factors, weights, and analyses used to generate risk scores and recommendations, with legal scholars suggesting that transparency is necessary on constitutional grounds. The PSA was developed through a process of unpublished research and the risk factors were initially not made available to the public, LJAF has since shared the factors and weights publicly.[11]

Third, judges are somewhat mixed on their perspective about racial bias during the pretrial phase and the role of the PSA. Some judges expressed that pretrial is an objective system devoid of bias, and they are not conscious of race, gender, or other ascribed statuses. However, given the findings from cognitive science and unconscious bias, one is left to wonder how such over reliance on the objectiveness of the system might inadvertently contribute to bias. Other judges were more aware of the subtleties of bias and were appreciative that the PSA does not include irrelevant factors such as homelessness, but instead focused on current and past behaviors. And, still others, challenged the reliance on criminal history items due to differential enforcement, prosecution, and sentencing. The nuanced views about racial disparity hopefully signals a willingness from the judiciary to be reflective about the potential for disparities on both conscious and unconscious levels. The U.S. has a complicated history of racial oppression and bias that includes well-known patterns of hyper-policing and enforcement of communities of color. Future research is needed to disentangle how the associations between criminal history factors may exacerbate the unfair treatment of people of color.

---

[11] The PSA risk factors, weights, and scales are available here: http://www.arnoldfoundation.org/wp-content/uploads/PSA-Risk-Factors-and-Formula.pdf

The PSA was developed as tool and not something to supplant judicial discretion. Recently, Berk et al. (2017) pointed out that the true utility of a risk assessment has to do with more than the robustness of any model to predict outcomes. Rather, more research is needed to know how assessments are implemented, understood, and used by practitioners, and stakeholders need to weigh-in to settle the many normative matters involved in risk assessments.

References

Angwin, J., Larson, J., Mattu, S., and Kirchner, L. (2016). *Machine bias: There's software used across the country to predict future criminals and it's biased against blacks.* Available online at www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (last accessed December 15, 2017).

Ares, C., Rankin, A., and Sturz, H. (1963). The Manhattan Bail Project: An interim report on the pretrial use of pretrial parole. *New York University Law Review, 38:* 67- 95.

Baradaran, S. and McIntyre, F. (2012). Predicting violence. *Texas Law Review, 90*: 496-570

Barbaree, H.E., Langton, C.M. and Peacock, E.J. (2006). Different Actuarial Risk Measures Produce Different Risk Rankings for Sexual Offenders. *Sex Abuse, 18*: 423.

Bechtel, K., Lowenkamp, C., and Holsinger, A. (2011). Identifying the Predictors of Pretrial Failure: A Meta-Analysis, *Federal Probation*, 75 (2).

Berk, R., Heidari, H., Jabbari, S., Kearns, M., and Roth, A. (2017). Fairness in criminal justice risk assessments: The state of the art. Working Paper found on January 5, 2018 here: https://arxiv.org/pdf/1703.09207.pdf

Bonta, J. (1996). Risk-needs assessment and treatment. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 18-32). Thousand Oaks, CA: Sage.

Campbell, M., French, S., & Gendreau, P. (2007). *Assessing the utility of risk assessment tools and personality measures in the prediction of violent recidivism for adult offenders.* Ottawa, Ontario, Canada: Department of Safety and Emergency Preparedness.

Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data, 5*(2): 153-163.

Danziger, S., Levav, K., and Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Science, 108*(17): 6889-6892.

Desmarais, S., Johnson, K., and Singh, J. (2016). Performance of recidivism risk assessment instruments in US correctional settings. *Psychological Services, 13*(3): 206-222.

Feeley, M.M., Simon, J. (1992). The New Penology: Notes on the Emerging Strategy of Corrections and Its Implications. *Criminology, 30:* 449.

Flores, A., Bechtel, K., and Lowenkamp, C. (2016). False positives, false negatives, and false analyses: A rejoinders to "Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks." *Federal Probation, 80*(2): 38-46.

Frank, Jerome. (1930). *Law and the Modern Mind.* New York: Brentano's.

Frase, R., Roberts, J., Hester, R., and Mitchell, K. (2015). *Criminal history enhancements sourcebook.* University of Minnesota Law School, Robina Institute of Criminal Law and Criminal Justice.

Frederick, S. (2002). Automated choice heuristics. pp. 548-558 in T. Gilovich, D. Griffin, and D. Kahneman [eds]. *Heuristics & Biases: The Psychology of Intuitive Judgment*. New York. Cambridge University Press.

Garfinkel, H. (1956). Conditions of Successful Degradation Ceremonies. *American Journal of Sociology*, 61(5): 420-424.

Garland, D. (2001*). Culture of Control: Crime and social order in contemporary society*. Chicago, IL: University of Chicago Press.

Gottfredson, S. D., and Moriarty, L. J. (2006) Statistical Risk Assessment: Old Problems and New Applications. *Crime & Delinquency, 52*(1): 178–200.

Gouldin, L. (2016). Disentangling flight risk from dangerousness. *Brigham Young University of Law, 837*(3): 838-898.

Grove, William M.; Meehl, Paul E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law*, 2(2): 293-323.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., and Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*: 19-30.

Guthrie, C., J.J. Rachlinski, and A.J. Wistrick. 2001. Inside the Judicial Mind. *Cornell Law Review 86:*777-830.

Guthrie, C., Rachlinski, J., and Wistrich, A. (2007). Blinking on the bench: How judges decide. *Cornell Law Faculty Publications*, Paper 917.

Hannah-Moffat, K. (2013). Actuarial sentencing: An unsettled proposition. *Justice Quarterly, 30*(2): 270-296.

Harcourt, Bernard E. (2010). Risk as a Proxy for Race. Criminology and Public Policy, *University of Chicago Law & Economics* Olin Working Paper No. 535; University of Chicago Public Law Working Paper No. 323.

Harris, Grant T., Rice, Marnie E., Quinsey, Vernon L., Lalumière, Martin L., Boer, Douglas, and Lang, Carol (2003). A Multisite Comparison of Actuarial Risk Instruments for Sex Offenders. *Psychological Assessment,* 15(3): 413-425

Hessick, C. and Hessick, F. (2011). Double Jeopardy as a Limit on Punishment, *97 Cornell L. Rev.* 45-86.

Hoschild, J. and Weaver, V. (2007). The skin color paradox and the American racial order. *Social Forces, 86*(2): 643-670.

Kahneman, D. (2011). *Thinking, fast and slow.* New York: Farrar, Straus and Giroux.

Kahneman, D. and Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. pp. 49-81 in T. Gilovich, D. Griffin, and D. Kahneman [eds]. Heuristics & Biases: The Psychology of Intuitive Judgment. New York. Cambridge University Press.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica, 47(2): 263–291.*

Kang, J., Bennett, M., Carbado, D., Casey, P., Dasgupta, N., Faigman, D., Godsil, R., Greenwald, A., Levinson, J., and Mnookin, J. (2012). Implicit bias in the courtroom. *UCLA Law Review, 59*: 1124-1185.

Kleinberg, J., Mullainathan, S., and Raghavan, M. (2016, forthcoming). Inherent trade-offs in the fair determination of risk scores. To appear in *Proceedings of Innovations in Theoretical Computer Science.* https://arxiv.org/abs/1609.05807

Lowenkamp, C., VanNostrand, M., and Holsinger, A., (2013). *Investigating the Impact of Pretrial Detention on Sentencing Outcomes*, Laura and John Arnold Foundation.

Mahon, T. (2012). *Justifying the Use of Previous Convictions as an Aggravating Factor at Sentencing*, 2012 Cork Online Law Review 85-87, http://corkonlinelawreview.com/index.php/2012/03/13/justifying-the-useof-previous-convictions-as-an-aggravating-factor-in-sentencing/.

Mamalian, C. (2011). *State of the science of pretrial risk assessment*. Pretrial Justice Institute.

Meehl, P. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.* Minneapolis: University of Minnesota

Miller, J. and Maloney, C. (2013). Practitioner compliance with risk/need assessment tools: A theoretical and empirical assessment. *Criminal Justice and Behavior, 40*(7): 716-736.

Minton, T. and Zeng, Z. (2016). *Jail inmates in 2015*. Washington, DC: US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.

Monahan, J., Skeem, J., and Lowenkamp, C. (2017). Gender, risk assessment, and sanctioning: The cost of treating women like men. *Law and Human Behavior, 40*, 580–593.

Papillon, K. (2013). Neuroscience and judicial decision making in criminal sentencing. *Court Review, 49*: 48-62.

Pretrial Justice Institute. (2015). *Issue Brief: Pretrial risk assessment: Science provides guidance on assessing defendants.* May, 2015.

Sacks, M. and Ackerman, A. (2012). Bail and sentencing: Does pretrial detention lead to harsher punishment. *Criminal Justice Policy Review, 25*(1): 59-77.

Schonteich, M. (2011). *Pretrial detention and health: Unintended consequences, deadly results: Literature review and recommendations for health professionals*. New York: Open Society Foundation.

Skeem, J. L. and Lowenkamp, C. T. (2016). Risk, race, & recidivism: Predictive bias and disparate impact. Social Science Research Network. *Criminology, 54*(4): 680-712.

Simon, D. (2004). A third view of the black box: Cognitive coherence in legal decision making, University of Chicago Law Review, 71(511): 511-586.

Stanovich, K.E., & West, R.F. (2002). Individual differences in reasoning: Implications for the rationality debate? In T. Gilovich, D. W. Griffin, D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 421-440). New York: Cambridge University Press.

Starr, S. (2014). Evidence-Based Sentencing and the Scientific Rationalization of Discrimination. *Stanford Law Review,* 66(4): 803-72.

Tonry, M. (2014). Legal and ethical issues in the prediction of recidivism. *Federal Sentencing Reporter, 26*(3): 167-176.

Tversky, A. and Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, 185*(4157), 1124-1131.

Vigione, J., Rudes, D., and Taxman, F. (2015). Misalignment in supervision: Implementing risk/needs assessments in probation. *Criminal Justice and Behavior, 42(*3): 263-285.