

THE ROLE OF EVIDENTIARY WEIGHT IN JUDICIAL SENTENCING
DETERMINATIONS

by ESTHER NIR

A Dissertation submitted to the

Graduate School-Newark

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Graduate Program in Criminal Justice

Written under the direction of Dr. Elizabeth Griffiths

and approved by

Newark, New Jersey
May 2016

ABSTRACT

The type and quantity of evidence in a case is a critical factor for deciding guilt, but should have little or no influence on the sentencing determinations of judges post conviction; this is because case evidence goes to guilt decisions by triers of fact, whereas sentences are imposed upon those already convicted. This dissertation examines the extralegal influence of evidentiary type and quantity on post conviction sentencing decisions using a mixed methods approach. Quantitative results demonstrate that violent felony trial cases with forensic evidence, and those cases with more varied pieces of physical evidence, result in longer custodial sentences for convicted defendants, whereas the presence of eyewitnesses fails to influence sentencing punitiveness. Qualitative interviews of 41 state court sentencing judges provide explanations for these relationships. The findings show that judges perceive cases with forensic evidence to be more objective and reliable than non-scientific evidence. While they appreciate the human quality of witness testimony, inherent credibility and reliability concerns result in decreased perceptions of evidentiary strength. Further, additional pieces of evidence corroborate other case evidence and improve perceptions of evidentiary weight. Thus, cases with more evidence and cases with forensic evidence increase judicial confidence levels in guilt. Finally, analyses utilizing a series of vignettes to evaluate sentencing rationales, as well as direct queries, reveal that judges impose longer sentences when their perceptions of evidentiary strength are highest *because* they are more confident in the guilt of the defendant.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the talent and dedication of my dissertation committee. My dissertation chair, Dr. Elizabeth Griffiths, exemplifies the term ‘mentor’ in every sense of the word. From the early conceptual stages, Beth always encouraged me to leave no stone unturned and to never accept mediocrity. She has guided me through every step of the process, generously sharing her wealth of knowledge, creativity, and incomparable ability to motivate and inspire. Beth is my advisor extraordinaire and I was fortunate to have her in my corner. I also am privileged to have Dr. Jody Miller as a member of my dissertation committee. Always generous with her time, Jody provided me with outstanding advice regarding how to approach prospective elite interviewees, conduct qualitative interviews, and analyze data. Jody’s practical, thorough, and straightforward strategies combined with her timely advice during critical stages of the data collection process enabled the collection of rich qualitative data necessary for the successful completion of this study. I also am lucky to have Dr. Rod Brunson serve on my dissertation committee. Rod provided excellent advice regarding how to conduct qualitative interviews and his feedback was instrumental in helping me improve my interview guide approach. Rod was particularly helpful in motivating me to continue recruitment efforts to increase judicial participation in this dissertation. I also am fortunate to have Dr. Brian Johnson serve as an outside reader. Brian’s expertise in sentencing scholarship and his ability to provide on point and practical suggestions enabled me to improve this dissertation. Brian’s suggestions during my prospectus defense encouraged me to consider methodological changes that strengthened the study.

This study would not have been possible without the gracious cooperation and participation of the 41 state criminal term judges who generously shared their perspectives from the bench. It was truly a privilege to hear their opinions, perceptions, and stories and an experience that I will always remember.

Finally, I would like to thank my parents, in-laws, sister, children, and a few special friends for their unwavering support. From helping me balance the rest of my life during my doctoral studies, to proofreading numerous pages of my dissertation, to providing me with honest and thoughtful opinions, to consistently encouraging my efforts, I am blessed with supportive and loving family and friends that are always behind me. Last, and most important, to my husband, Dror: There are no words that can adequately express how your endless help, love, and support for my every endeavor has inspired me to accomplish seemingly unreachable goals. Let it suffice to say that this dissertation is truly for you.

Table of Contents

Chapter 1. Introduction.....	1
Dissertation Outline.....	5
Chapter 2: Literature Review.....	9
Judicial Discretion in Sentencing.....	9
Sentencing Factors.....	25
The Role of Evidentiary Weight at Sentencing.....	40
Chapter 3. Methodology.....	60
Quantitative Analyses.....	61
Qualitative Analyses.....	67
Interview Guide.....	72
Vignettes.....	76
Benefits and Limitations.....	81
Chapter 4. Estimating Evidentiary Weight on Sentence Length.....	86
Description of Variables.....	86
Results.....	92
Discussion.....	95
Chapter 5. Judicial Perceptions of Evidentiary Weight.....	101
Forensic Evidence.....	103
DNA.....	103
Fingerprints.....	110
Ballistics.....	114
Chemists.....	117

The Need for Forensic Evidence.....	120
Witness-based Evidence.....	122
Eyewitnesses.....	123
Character Witnesses.....	126
Alibi Witnesses.....	128
Psychologists.....	130
Comparisons of Strength between Evidentiary Type.....	140
Conclusion.....	142
Chapter 6. The Influence of Evidentiary Weight at Sentencing Post-Verdict.....	146
Direct Inquiry.....	150
Vignettes.....	156
Scenario A.....	161
Scenario B.....	165
Scenario C.....	167
Scenario D.....	171
Comparisons between Scenarios A & B.....	175
Comparisons between Scenarios C & D.....	177
Conclusion.....	180
Chapter 7. Conclusion.....	183
References.....	199
Appendix 3.1. Informed Consent Form.....	227
Appendix 3.2. Interview Guide.....	232
Appendix 3.3. Vignettes.....	243

List of Tables

Table 3.1. Judicial Characteristics.....	217
Table 4.1. Descriptive Statistics.....	218
Table 4.2. Evidentiary Weight by Crime Type.....	219
Table 4.3. Ordinary Least Squares Regression Predicting Sentence Length for Defendants Convicted of Violent Crimes.....	220
Table 4.4. Ordinary Least Squares Regression Predicting Sentence Length for Defendants Convicted of Violent Crimes at Trial.....	221
Table 4.5. Ordinary Least Squares Regression Predicting Sentence Length for Defendants Convicted of Violent Crimes via Pleas.....	222
Table 6.1. Description of Vignettes.....	223
Table 6.2. Sentences Imposed by Judges in the Vignettes.....	224

List of Figures

Figure 4.1. Distribution of Sentence Length.....	225
Figure 4.2. Distribution of Logged Sentence Length.....	226

CHAPTER 1. INTRODUCTION

The challenge of achieving justice in an era of mass incarceration is not only one of quantity – reflected in the excessive number of people incarcerated per capita relative to other nations – but also of quality. That is, scholars have repeatedly demonstrated significant disparities in arrest, conviction, disposition, and sentencing by race, ethnicity, gender, and other socioeconomic and demographic characteristics (Albonetti, 1997; Doerner & Demuth, 2010; Everett & Wojewicz, 2002; Feldmeyer & Ulmer, 2011; Steffensmeir & Demuth, 2000). The existence of these disparities is suggestive of sizable substantive injustices in the way that the criminal justice system operates.

It is well established that, holding the legally relevant features of criminal cases constant, extralegal factors have some effect on case outcomes at various stages of criminal justice processing (Albonetti, 1997; Mustard, 2001; Rodriguez et al., 2006; Steffensmeir & Demuth, 2000); yet, little attention has been focused on the evidentiary characteristics of cases as an extralegal consideration. One explanation for this lack of focus is that the type (e.g. forensic or witness-based), and strength (e.g. direct or circumstantial, quantity of evidence) of evidence – or the evidentiary weight – of cases is assumed to be legally relevant to explaining all aspects of case outcomes, even though it is not. For example, the type and strength of evidence presented during the guilt phase of a criminal prosecution is clearly imperative to the trier of fact in reaching a determination of guilt or innocence. For this reason, police and prosecutors expend considerable effort to collect and present credible and compelling evidence to bolster their cases against criminally charged defendants during the guilt phase of a criminal case. Evidentiary weight loses its legal relevance at sentencing, however, for reasons described below.

Nonetheless, what little research exists on this topic suggests that sentences are improperly influenced by evidentiary weight (Peterson et al., 1987; Peterson et al., 2010; Spohn, 2000).

The rationale for considering evidentiary weight as an extralegal sentencing characteristic stems from the fact that once a legally designated trier of fact – in accordance with existing statutory and constitutional protections and standards – determines that the defendant is guilty of the crime as charged, it is beyond the role of the sentencing judge to re-evaluate the evidence used to convict the defendant. To the contrary, a defendant convicted on the basis of a “weaker” evidentiary package is no less guilty in the eyes of the law than a defendant convicted on the basis of a “stronger” or more scientific package. While a trial judge can overturn a jury verdict in situations where he or she determines that no reasonable trier of fact could have found the defendant guilty, the question of whether the defendant is guilty has already been resolved once the verdict is upheld and the case moves to the sentencing phase. A judge who is influenced by guilt phase evidentiary weight during sentencing is, in essence, usurping the role of the jury by infusing his or her own personal perceptions of the strength of the case into his or her sentence determination.¹ Consequently, a judge’s consideration – even if inadvertent – of evidentiary weight at sentencing may lead to disparate sentences among similarly situated defendants. These disparities, in turn, undermine the quality of justice achieved within the American criminal justice system.

This dissertation systematically examines whether and how judges use evidentiary weight as a factor in reaching their sentencing decisions using a mixed methods

¹ See pages 52-54 for a more complete analysis of, and justification for, categorizing judicial consideration of evidentiary weight as extralegal at sentencing.

approach. The combination of quantitative and qualitative methods to address questions about the role of evidentiary factors for judicial sentencing decisions is ideal for several reasons. First, in order to examine the associations between various quantities and types of evidence (i.e. forensic, witness based) and the severity of sentences imposed upon convicted defendants, a quantitative approach allows for an examination of *whether* and *to what extent* evidentiary weight acts as an extralegal sentencing factor that contributes to sentencing disparities among hundreds of criminal cases. To accomplish this, I utilized a National Institute of Justice (NIJ) dataset containing a large variety of evidentiary variables (forensic and witness-based) and case disposition information (including sentence length) to examine the presence and strength of these associations.

These quantitative analyses cannot elucidate, however, *how* or *why* any such pattern arises. Thus, inferences about how these factors operate to produce sentencing disparities are speculative on the basis of the quantitative analyses alone. Said differently, the pattern of quantitative results may be suggestive of the process linking evidentiary weight to disparities in sentencing but it is, of necessity, tentative. For example, one could argue that a finding in which the introduction of forensic evidence at trial is associated with longer prison terms at sentencing suggests that judges possess greater confidence in a defendant's guilt when the direct evidence is perceived as scientifically reliable. While logical, this explanation involves a speculative leap that goes beyond the scope of the results.

Since judicial perceptions on sentencing are required to confirm these speculative explanations, first-hand interviews with sentencing judges were conducted to provide unique insight that is only available through direct dialogue with sentencing decision

makers. In total, I conducted 41 in-depth interviews with state court judges, from 18 counties throughout New York State, who are empowered to try and sentence defendants on violent felony offenses. These interviews explored, among other issues, the specific decision making processes and rationales judges use in sentencing, with an emphasis on understanding how evidentiary weight influences sentencing outcomes in criminal cases.

My project explores a series of research questions related to the type and quantity of evidence presented in criminal cases using the NIJ dataset. Foremost among these is: Are cases characterized by greater evidentiary weight associated with longer prison sentences for convicted defendants? Specifically, I examine how the existence of different pieces of evidence presented during the guilt phase of a criminal case influences sentence severity for trial and plea cases (e.g. is forensic evidence, typically perceived as a more objective and reliable form of evidence, associated with longer sentences?) I also examine whether and how the quantity of physical evidence in a case influences sentence length.

While these analyses demonstrate that evidentiary type and quantity influence sentence severity post-conviction, the influence of evidentiary weight on *judicial* sentencing decisions cannot be fully explored in analyses utilizing plea cases; this is because sentencing in plea cases are not a pure reflection of judicial discretion (see, Johnson, 2003). In plea cases, the guilt and sentencing stages usually occur concurrently (i.e. a defendant's decision to accept a plea bargain is conditioned upon a promised sentence) and are largely a reflection of prosecutorial discretion. While a judge can refuse to accept a plea, which he or she deems to be inappropriate, judges are usually presented with (and impose) pre-determined, negotiated sentences (i.e. negotiated between the

prosecution and the defense). In contrast, sentencing post-verdict is a separate stage of case processing and a purer reflection of judicial discretion (Johnson, 2003). Once a verdict is rendered, it is within the purview of the sentencing judge (although judges will consider recommendations from other criminal justice actors) to decide the appropriate punishment. For these reasons, I also examine trial cases exclusively. These cases provide an avenue to explore judicial decision making in an arena of considerable judicial discretion.

Based on the in-depth qualitative interviews with sentencing judges, I then focus on judicial perceptions of various evidentiary forms and explore whether judges explicitly base their confidence in the defendant's guilt on aspects of evidentiary weight. Finally, using a series of vignettes and interview questions I examine whether judges consider evidentiary weight in the severity of the sentences they hand down for convicted defendants. These vignettes are based on post-trial jury convictions, where judicial exposure to case evidence and sentencing discretion is greatest.

Outline of the Dissertation

This dissertation consists of the following chapters: (2) Literature Review; (3) Methodology; (4) Estimating Evidentiary Weight on Sentence Length; (5) Judicial Perceptions of Evidentiary Forms; (6) The Influence of Evidentiary Weight at Sentencing and (7) Conclusion.

In Chapter 2, I review the literature relevant to this study and describe how this dissertation contributes to prior scholarly efforts. I begin by examining earlier work relating to judicial sentencing discretion; indeed, any study exploring how and why judges act in a particular manner necessitates an understanding of the scope of judicial

discretionary powers, knowledge of the ways those powers are exercised, and existing challenges. Following a review of this literature, I examine the legally relevant factors found to influence sentence severity including seriousness of the offense, the defendant's criminal history and plea bargaining practices. I also explore extralegal influences historically linked to sentencing disparities including defendant and victim demographic characteristics (i.e. race/ethnicity, gender, age) and the locality of the sentencing court.

Following this review, I shift focus to an area that has received little scholarly attention: the role of evidentiary weight on judicial sentencing decisions. To establish the necessary foundations, I first define evidentiary weight and review scholarly studies pertaining to the strengths and weaknesses of different forms of evidence (i.e. forensic and witness-based). Relying on case law and other argument, I support my position that evidentiary weight should be viewed as an extralegal sentencing influence, particularly in trial cases. I conclude Chapter 2 by reviewing the small body of literature that examines the influence of various types of evidence on sentence severity.

Chapter 3 outlines the mixed methods employed in this study. I begin by describing the quantitative methodology, including the evidentiary and control variables, and OLS regression models utilized in these analyses. I then discuss the qualitative methods employed. I describe the sample of judges, recruitment methods, and interview strategies. The structure of the interviews are outlined in Chapter 3. The first component consists of an interview guide approach, with a focus on sentencing discretion, sentencing factors, judicial perceptions of various forms of evidence, judicial perceptions of juries and the influence of evidentiary weight on sentencing decisions. I conclude the methodology chapter by describing the series of four vignettes employed during the

interviews to extract the thought processes engaged by judges in determining sentences, not evident by direct inquiry alone. I describe the construction of the vignettes, and my rationale for including or excluding specific case information.

In Chapter 4, I analyze the influence of type and quantity of evidence on sentence severity among hundreds of violent felony cases. Regression results of two models are discussed; one model tests the influence of forensic evidence, eyewitnesses and the quantity of physical evidence on both trial and plea convictions and the other model tests the model among trial cases exclusively. Both models establish that evidentiary weight influences sentencing decisions. The nature and implications of these influences are discussed in Chapter 4. The analyses reported in Chapter 4 are forthcoming in the *Criminal Justice Policy Review* (Nir & Griffiths, forthcoming).

Chapter 5 is the first qualitative analytic chapter of this dissertation. In it, I explore judicial perceptions of various forms of forensic and witness-based evidence. Forensic types analyzed include DNA evidence, fingerprint analyses, ballistics evidence and chemist testimony. Witness evidence includes eyewitnesses, character witnesses, alibi witnesses and psychologists. I describe the processes by which judges evaluate witness credibility and reliability and how these evaluations influence perceptions of evidentiary strength. Finally, I compare judicial perceptions of forensic and witness-based evidence; judges perceive forensic evidence as a stronger and more objective form of evidence than non-scientific witness testimony. Chapter 5 concludes by demonstrating that, for many judges, forensic evidence increases judicial confidence in a defendant's guilt.

In Chapter 6, I build on prior findings and analyze how evidentiary weight influences sentencing decisions. The first part of the chapter reports judicial responses to direct questions regarding the influence of evidentiary weight on sentencing decisions (i.e. Do you consider the strength of the evidence in determining sentence? In what way do your own perceptions of guilt and/or the accuracy of the jury's verdict factor into your sentencing decisions?). The discussion then shifts to analyses of judicial reactions to four post-trial hypothetical case scenarios. Judicial comments relating to these vignettes, as well as hypothetical sentence lengths imposed, are evaluated to explore sentencing processes that may be subconscious or, at least, not apparent from other interview discussions.

In Chapter 7, I draw from the quantitative and qualitative findings and conclude that evidentiary weight – both type and quantity of evidence – has a profound influence on judicial sentencing decisions post-conviction. These findings are then placed within the context of the broader literature regarding extralegal sentencing factors. These conclusions are important because judicial consideration of evidentiary weight post-conviction leads to unjust and inequitable sentences for convicted defendants, despite receiving very little attention in the literature. I conclude by discussing this study's limitations and important future research directions.

CHAPTER 2. LITERATURE REVIEW

In this chapter, I review the theoretical foundations critical to assessing the role of evidentiary weight in judicial sentencing decisions. I begin by exploring the complexities of judicial sentencing discretion. Developing an understanding of how judges exercise their discretion is an important first step in efforts to control the problem of sentencing disparities. I then review the literature regarding legally relevant (e.g. seriousness of the offense and criminal history) and extralegal (e.g. defendant and victim demographic characteristics) sentencing factors and their influence on judicial decision making. In the final section of this chapter, I explore the relative strengths and weaknesses of different types of evidence and the extralegal role evidentiary weight plays during sentencing.

A. Judicial Discretion in Sentencing

The sentencing phase of a criminal case is a critical event in the lives of convicted defendants. Decisions made by sentencing judges can result not only in a loss of freedom, but also can dramatically impact the lives and fortunes of defendants' families and communities (Clear, 2008). Unfortunately, history shows that judicial sentencing determinations are fraught with unjust disparities. Much of the blame for these disparities in sentencing for like-offenses has been attributed to the considerable discretionary power wielded by sentencing judges (Frankel, 1972; Lynch, 2009).

As a general rule, judges are afforded a tremendous amount of discretion at sentencing. Consequently, much scholarly attention focuses on the scope, complexities, benefits, and problems inherent in this process (Johnson, 2005; Johnson & DiPietro, 2012; Savelsberg, 1992; Steffensmeir, Ulmer & Kramer, 1998; Ulmer & Kramer, 1996). In this section, I address several issues related to judicial sentencing discretion. First, I

provide a working definition of sentencing discretion and discuss the areas (e.g. wide sentencing ranges, legally sanctioned departures from sentencing guidelines) in which it emerges. Next, I highlight some of the problems involved in granting these broad discretionary powers to judges. Finally, I describe the history of various legislative efforts enacted to control sentencing discretion in past decades, outline New York State's current sentencing scheme and the current federal sentencing guidelines² and discuss the ability of legislative efforts to reduce disparities in the aftermath of *Booker*.

Judicial Discretion Defined

Judicial sentencing discretion refers to the freedom of judges to make sentencing determinations within legally sanctioned ranges established by the legislature (see Bushway & Forst, 2013). The ability to exercise discretion among fairly wide ranges provides judges with enough flexibility to make individualized decisions based on the unique circumstances of a given case. While judicial sentencing discretion is the target of a great deal of criticism, “discretion is inherently neither good nor bad. It can be used skillfully to counter ill-conceived or vague laws and policies . . . but it can be misused as well, with immediate harm to the victims of the abuse and long-lasting harm to the legitimacy of our system of criminal justice” (Bushway & Forst, 2013, p. 215). While a certain degree of discretion is necessary for judges to be able to effectively handle the

² The implementation of the Federal Sentencing Guidelines marked a concerted effort to control judicial sentencing discretion. Consequently, it is an ideal example (and therefore highlighted in this section) to use in studying whether and how judicial discretion at sentencing can be effectively controlled by legislative measures. Indeed, much scholarly attention has been focused on examining the influence of the Federal Sentencing Guidelines. Moreover, the shift in status from mandatory to advisory in 2005 (see *U.S. v. Booker*, 2005) provides a unique opportunity to compare its influence on sentencing decisions in different legislative environments. While the jurisdictions studied in this dissertation are state jurisdictions (i.e. the qualitative data collection was conducted in New York State and the quantitative analyses utilized data from Indiana and California), lessons learned from the extensive studies conducted on the influence of the Federal Guidelines, both pre and post *Booker*, provide an important perspective uniquely available from the federal arena that can inform our understanding of the decisions of state court judges as well.

unique circumstances of every case, determining how much discretion any legal actor should have is worthy of serious consideration and debate. It is unclear “where boundaries of law should end and where officials should begin to exercise discretion in interpreting and enforcing the laws” to ensure fair and just outcomes for defendants convicted of similar offenses (Bushway & Forst, 2013, p. 200).

Discretion is “made available” to legal actors, including judges, “by a law or rule” (Bushway & Forst, 2013, p. 200; see also Johnson, 2005), and can take a variety of different forms. For example, Bushway and Forst (2013) classify discretion into two distinct categories distinguished by actor: these are Type A and Type B forms of discretion. Type A refers to “the discretion that individual actors [such as judges or prosecutors, for example] have to make decisions with variation given a set of rules” and Type B refers to “discretion that legislators and criminal justice policy-makers have to establish a set of rules” (Bushway & Forst, 2013, p. 201). The establishment of sentencing guidelines by a federal or state legislative body is an example of Type B discretion. In contrast, the imposition of legally sanctioned sentences by judges, within a sentencing guidelines range, is an example of Type A discretion. Accordingly, a judge who sentences a defendant to a period of ten years, when the sentencing range is between five and fifteen years, is exercising Type A discretion; he or she is working within the legally proscribed framework.

Conceptualizing Discretion

To date, scholars have tended to distinguish formally rational standards from substantively rational considerations. Specifically, formal rational standards are employed when judges make sentencing decisions on the basis of legal rules (Johnson,

2005). For example, a judge who sentences a defendant according to a sentencing guidelines scheme is sentencing according to formal rational standards. In contrast, substantive rationality “entails consideration of defendants’ particularistic circumstances, needs or characteristics as well as the practical consequences of sentences for individuals and organizations” (Ulmer & Kramer, 1996, p. 384; see also Johnson, 2005; Savelsberg, 1992). Focusing on guideline departures (which reflect an area of greater discretionary freedom),³ Johnson contends that “judicial departures can be understood as the result of the complex interplay between formally rational guideline recommendations and substantively rational sentencing concerns, based on varying interpretations of different focal concerns across courtroom communities” (Johnson, 2005, p. 768).

Substantively rational sentencing concerns may involve consideration of a variety of factors that local judges believe are important sentencing criteria, based on specific “focal concerns.” According to a focal concerns perspective, judges reach decisions based on three primary concerns: blameworthiness of the defendant, protection of the community, and practical concerns (Kramer & Ulmer, 1996; Steffensmeier & Demuth, 2006; Steffensmeier, Ulmer & Kramer, 1998; Ulmer & Johnson, 2004). Blameworthiness – generally associated with a retributivist punishment philosophy – focuses on the defendant’s actions and the harm caused as a result of those actions. Perceptions of blameworthiness may be based on factors such as the seriousness of a given offense,

³ In order to effectively study how discretion is exercised, it is helpful to focus on the areas where sentencing judges possess discretionary powers. For example, judges in many jurisdictions work within specified sentencing guideline schemes that limit their flexibility to reach individualized decisions during sentencing. Yet, sentencing guidelines do not remove all discretion. There are areas where judges, legally abiding by set guideline rules, still retain the opportunity to exercise their own discretion. For example, upward and downward departures within sentencing guidelines schemes provide judges with the opportunity to insert their own concerns and priorities into the sentencing process (Johnson, 2005).

including the resulting harm and the criminal history of the defendant. Protection of the community – a second focal concern – is associated with an incapacitation goal of punishment and focuses on predicting the future dangerousness of the defendant (i.e. the likelihood of recidivism). Like blameworthiness, these predictions may be based on factors such as the defendant’s criminal history, details of the offense, and specific defendant characteristics. Finally, practical concerns may involve both organizational and individual factors. Organizational factors may include caseload pressures and concerns regarding prison overcrowding in the jurisdiction while individual concerns take specific offender needs – such as an offender’s responsibility for dependent children – into account. These focal concerns may be shaped, in practice, by the defendant’s social structural position (e.g. race/gender) and by other political or public considerations (Steffensmeier & Demuth, 2006; see also Nowacki, 2015).

Kramer and Ulmer (1996) test these theoretical perspectives in an examination of the sentencing criteria used by judges in Pennsylvania who deviated from the Pennsylvania Sentencing Guidelines. They argue that judges rely on a variety of information in reaching determinations about whether they should depart from the guidelines, including the severity of the current offense, the defendant’s criminal history, race, gender, and plea agreements; said differently, judges use these factors to stereotype defendants. Stereotypes enable judges to “project behavioral expectations” regarding blameworthiness, whether the defendant is a danger to the community, and whether the defendant is a good or bad rehabilitation risk (Kramer & Ulmer, 1996, p. 98).

The Problem of Judicial Discretion

While judicial discretion is a necessary component of the criminal justice sentencing process, there are inherent problems associated with vesting such discretion in sentencing judges. Specifically, sentencing discretion, by definition, involves giving judges some degree of freedom to reach their own individualized and distinct decisions. This freedom, even when skillfully exercised with the best of intentions, will naturally result in sentencing variations – where defendants convicted of similar offenses and possessing similar criminal histories receive disparate sentences. These disparities are particularly problematic when the sentences include periods of incarceration. Such inequities can generate perceptions that the system is unjust, which serve to undermine system legitimacy (Frankel, 1972; Lynch, 2007-2008; Lynch, 2009).

Critics of the U.S. sentencing system argue that personal differences among judges exacerbate the problem of sentencing disparities. In his book *Law without Order* (1972), for example, Judge Marvin Frankel suggests that differences in judicial backgrounds, personalities, and life experiences interfere with the equitable imposition of criminal sentences. According to Frankel (1972, p.7),

judges vary widely in their explicit views and “principles” affecting sentencing; they vary, too, in the accidents of birth and biography generating the guilts, the fears, and the rages that affect almost all of us at times and in ways we often cannot know.⁴

Some judges are perceived to be “tough” in sentencing defendants while other judges are perceived to be “lenient.” While some judges believe in the punishment theory of retribution, others see rehabilitation, incapacitation, or other rationales as the central goals of punishment (Forst & Rhodes, 1982). Variations in judicial philosophies

⁴ Judge Marvin Frankel’s words were a source of inspiration for the enactment of the Federal Sentencing Guidelines, created in part as a means to limit judicial discretion (Gertner, 1999).

regarding punishment goals (e.g retribution, incapacitation) can influence the manner in which judicial sentencing discretion is exercised and consequently, result in sentencing disparities (Bushway & Forst, 2013; Forst & Wellford, 1981). For example, a judge who favors a rehabilitation philosophy is more likely to sentence offenders to alternative rehabilitative treatment options instead of prison compared to a judge who approaches punishment with a just deserts philosophy (Lynch, 2007-2008). Forst and Rhodes (1982) identify five distinguishable sentencing goals among the 264 federal judges they interviewed; these goals include general deterrence, special deterrence, incapacitation, rehabilitation, and just deserts. Judges were asked to rate the level of importance of each goal on a five point scale ranging from extremely important to not important at all. While a quarter of the judges stated that they believed that rehabilitation was an important sentencing goal, almost as many (19%) felt that it was no more than slightly important. And one quarter of the judges believed that just deserts was an important sentencing goal, whereas nearly half (45%) articulated that it was only slightly important or not important at all (Forst & Rhodes, 1982).

Perceptions regarding the level of seriousness of a given offense may also vary among individuals (Ramchand et al., 2009; Samuel & Moulds, 1986; Wolfgang, 1985). For example, researchers in California conclude that individuals who are victims of crimes view offenses as being more serious than do non-victims (Samuel & Moulds, 1986). Judges also vary in whether they perceive characteristics of a defendant's background to be aggravating or mitigating. For example, judges vary in "whether they regard factors such as poverty or drug addiction as extenuating circumstances or markers of likely recidivism" (Lynch, 2009, p. 236).

In addition to differences stemming from the personal characteristics of judges, differences may arise as a function of organizational structure. That is, judges receive little or no training before assuming the role. As Frankel (1972, pp. 6-7) notes,

trial judges are almost totally unencumbered by learning or experience relevant to sentencing. Nothing they learned in law school touched on our subject more than remotely. The new judge may be discovered within days or weeks fashioning judgments of imprisonment for long years. No training, formal or informal, preceded the first of these awesome pronouncements.

While legislative measures provide some guidance, the sentencing structure in most jurisdictions provides untrained judges with tremendous discretion in reaching sentencing determinations (see, for example, Lynch, 2009).

Legislative Attempts to Limit Judicial Sentencing Discretion

In order to mitigate some of the problems caused by judicial sentencing discretion (e.g. the influence of race) and achieve “reasonable uniformity in sentencing” (U.S. Sentencing Commission, 2011, p. 2),⁵ various legislative measures have been adopted to control the amount and type of discretion judges are afforded during the sentencing phase of criminal trials (e.g. Minnesota Sentencing Guidelines, 2013; North Carolina Sentencing Guidelines, 2013; U.S. Sentencing Commission, 2013). Much legislation in this area is focused on providing structure to sentencing practice (Albonetti, 2011; Feeney, 2003; U.S. Sentencing Commission, 2011) in order to mitigate unjust sentencing disparities (Ulmer, Light & Kramer, 2011; U.S. Sentencing Commission, 2011).

While some jurisdictions have enacted sentencing guidelines, others rely on criminal law and procedure statutes that designate the punishment that judges may legally

⁵ Cites that reference U.S. Sentencing Commission, 2011 are referring to the 2011 Federal Sentencing Guidelines Manual.

impose for particular offenses to structure sentencing decisions. The structure of sentencing statutes varies by jurisdiction. For example, in New York State, the Penal Law and Criminal Procedure Law provide a sentencing structure within which a sentencing judge must operate (see, for example, N.Y.S. Penal Law, 2015). In this case, sentencing statutes provide for either determinate or indeterminate sentences based on various factors including whether or not the crime is classified as violent or non-violent and the defendant's criminal history (N.Y.S. Penal Law, 2015). Further, New York State classifies defendants into categories according to their prior convictions. For example, some classifications include "violent felony offender," "second violent felony offender," "persistent felony offender," and "persistent violent felony offender," among others.

Even within one jurisdiction like New York, sentencing ranges for certain crimes are very broad while others are much narrower. For example, the punishment range for a first offender on a Class B violent felony is a determinate sentence between 5 and 25 years. In contrast, the range for a first offender on a Class E violent felony is a determinate sentence between 1 1/2 and 4 years. The range for a Class E violent felony narrows to a determinate sentence between 3 and 4 years when the defendant is a violent predicate offender (N.Y.S Penal Law, § 70). Consequently, the amount of discretion afforded to New York State sentencing judges varies based on the crime type and the defendant's criminal background.

A review of sentencing statutes and guidelines and their influence on judicial sentencing discretion is helpful in illuminating the extent to which discretionary power leads to sentencing disparities. For example, jurisdictions adopting tighter legislative measures provide counterfactuals to those environments where judges are more

“liberated” to exercise their own discretion in enabling a comparison of fluctuations in disparity rates that occur under different conditions (see Ulmer, Light & Kramer, 2011). Alternatively, scholars can examine changes in sentencing determinations pre- and post-legislative efforts to limit discretion. Perhaps the most highly visible legislative measure to reduce judicial discretion at sentencing was the enactment of the Federal Sentencing Guidelines in 1987. Originally imposed as mandatory and subsequently relegated to an advisory status (as a result of the Supreme Court decision in *U.S. v. Booker*, 2005), the Federal Sentencing Guidelines provide a unique avenue to explore the influence of legislative measures on sentencing discretion over time and in different environments.

Prior to the enactment of the Federal Sentencing Guidelines, federal judges sentenced defendants based on indeterminate sentencing schemes that, by design, enabled and facilitated the use of generous degrees of latitude in judicial decision-making. For example, depending on the crime, defendants with similar criminal histories and convicted of similar offenses could be legally sentenced to punishments ranging from probation to decades in prison (Feeney, 2003). Individual judges were tasked with determining a sentence, constrained only by wide-ranging parameters and their personal judgments (Feeney, 2003; Frankel, 1972). Prior to the enactment of the guidelines, constitutional issues, occasional minimum sentences, and the maximum statutory penalty for a given offense were the only legal restrictions placed upon federal sentencing judges. Moreover, judges were generally not required to provide explanations for their sentences, and their sentences were not usually subject to appellate review (Newman, 2002).

In 1987, when the Federal Sentencing Guidelines were enacted, adherence to the standards set by the guidelines became a mandatory requirement for federal judges. The

imposition of this detailed and mandatory sentencing scheme severely curtailed the exercise of judicial discretion among federal judges. As a result, studies could be conducted to compare sentencing disparities both prior and subsequent to the enactment of the guidelines to ascertain whether and how limiting the use of judicial discretion influences case outcomes. In 2005, however, in response to constitutional challenges, the Supreme Court rendered the federal sentencing guidelines advisory, rather than mandatory (*U.S. v. Booker*, 2005). Essentially, a significant degree of judicial discretion was thus returned to the hands of federal sentencing judges. The shift from mandatory to advisory status in 2005 provided an additional opportunity for scholars to explore the influence of added discretion on sentencing outcomes. Given the unique structure of these well-studied guidelines and their ability to inform our understanding of the influence of judicial discretion on sentencing, the next section will review the federal sentencing guidelines and some of the studies evaluating their impact on sentencing disparities.

Federal Sentencing Guidelines

In response to congressional concerns over disparities in sentencing (and critiques of the sentencing process by court actors such as Judge Marvin Frankel), Congress passed the Sentencing Reform Act of 1984 (Feeney, 2003). The Sentencing Reform Act established the U.S. Sentencing Commission. Under the Sentencing Commission, the Federal Sentencing Guidelines were created with the goal of achieving nationally uniform sentencing policies for criminal defendants. The Federal Sentencing Guidelines provide a sentencing table that directs judges to appropriate sentences in specific cases. The table creates a grid that cross-tabulates information on both “offense level” and a

defendant's "criminal history"⁶ (Feeney, 2003; U.S. Sentencing Commission, 2011). A sentencing range is specified at the point where offense level and criminal history intersect (Feeney, 2003).

An examination of the Federal Sentencing Guidelines illustrates the complexities involved in establishing a sentencing structure capable of equitably determining the proper sentence to impose upon a defendant based on the "nature and characteristics" of a given offense (18 U.S.C. Section 3553). According to the Federal Sentencing Guidelines Manual (2011, p. 4),

a sentencing system tailored to fit every conceivable wrinkle of each case would quickly become unworkable and seriously compromise the certainty of punishment and its deterrent effect. The list of potentially relevant features of criminal behavior is long; the fact that they can occur in multiple combinations means that the list of possible permutations of factors is virtually endless. The appropriate relationships among these different factors are exceedingly difficult to establish, for they are often context specific.

In establishing a sentencing structure in which offense seriousness is a major component, the federal guidelines waiver between creating a system that includes numerous and complex subcategories of offenses that would render the system "less workable" and a broad sentencing range that would risk "correspondingly broad disparity in sentencing, for different courts may exercise their discretionary powers in different ways" (U.S. Sentencing Commission, 2011, p. 4).⁷ Nonetheless, the Federal Guidelines are generated from data that evaluate pre-guideline sentencing practice (U.S. Sentencing Commission, 2011). While these guidelines are lengthy and contain most major distinctions that are relevant to sentencing determinations, other important distinctions –

⁶ For a complete discussion of both legal and extralegal sentencing factors, including the rationale for including criminal history as a legally relevant sentencing factor, please see Section B, *supra*.

⁷ One of the major criticisms of the guidelines is that it is "extraordinarily rigid, detailed, and cumbersome" (Newman, 2002, p. 320).

such as aggravating or mitigating circumstances not enumerated in the guidelines – may be addressed by sentencing judges by departing from the guidelines (U.S. Sentencing Commission, 2011). Departures from the guidelines are subject to appellate review and are restricted to certain circumstances. For example, judges are permitted (among several other delineated reasons) to make upward departures based on “unidentified circumstances” in the case that are “not adequately taken into consideration by the guidelines” (e.g. several victims were injured during the course of a robbery) (U.S. Sentencing Commission, 2011, Sec. 5K2.0). An example of a permissible downward departure is a case where the defendant has two or more characteristics or circumstances – each independently insufficient grounds for departure – that render the case exceptional if considered together (see U.S. Sentencing Commission, 2011, Sec. 5K2.0).

In addition to offense level, the Federal Sentencing Guidelines provide a working example of how criminal histories may be useful in reaching reasonably equitable sentencing determinations; in particular, the Federal Sentencing Guidelines establish an elaborate system of six criminal history categories to be used in calculating a defendant’s sentence. The categories determine the number of criminal history points assigned to a defendant based, in part, on prior sentences of imprisonment and convictions (U.S. Sentencing Commission, 2011).

Constitutional Challenges

In 2005, the constitutionality of the Federal Sentencing Guidelines was challenged in the case of *U.S. v. Booker, 2005*. In that case, the federal district court judge made an upward departure from the range identified in the guidelines and enhanced Booker’s prison sentence by almost ten years. The judge determined this sentence after

finding, by a preponderance of the evidence,⁸ that Booker possessed over 500 additional grams of crack cocaine. Booker's jury never heard the evidence regarding the possession of the additional drugs. Booker challenged the constitutionality of his sentence on Sixth Amendment grounds, arguing that his right to a trial by jury was violated since the judge enhanced his sentence using facts that were not reviewed by a jury. The Supreme Court held in favor of the defendant and ruled that a judge cannot enhance a defendant's sentence using evidence that was not presented at trial. As a result of this ruling, the court determined that holding the Federal Sentencing Guidelines to be mandatory (18 U.S.C.A. sec. 3553(b)(1); *U.S. v. Booker*, [2005]) was incompatible with its jury trial holding in *Booker*. Consequently, the Court determined that, to maintain constitutionality, the guidelines must be advisory instead of mandatory.⁹

Impact of Federal Sentencing Guidelines

Since the enactment of the Federal Sentencing Guidelines, numerous studies have evaluated whether the guidelines mitigate the problem of sentencing disparities (Everett & Wojewicz, 2002; Feldmeyer & Ulmer, 2011; Johnson & Betsinger, 2009; Steffensmeier & Demuth, 2006; Ulmer, Eisenstein & Johnson, 2010). Under mandate, the U.S. Sentencing Commission conducted some of these studies to determine whether the degree of sentencing disparities (i.e. range of sentences for similarly situated defendants) has lessened as a result of the implementation of the sentencing guidelines (U.S. Sentencing Commission Report, 1991). In a 1991 report, the Commission found that

⁸ Preponderance of the evidence is a significantly lower burden of proof than the standard of proof beyond a reasonable doubt necessary to obtain a guilty verdict in front of a jury.

⁹ See, also, *U.S. v. Fanfan* (2004) that was joined with *Booker*. See, also, *Gall v. U.S.* (2007), in which the Supreme Court likewise addressed constitutional issues concerning the Federal Sentencing Guidelines.

sentencing disparity rates between comparable offenders were substantially lower in a variety of offense types (e.g. bank robberies). In contrast, other research evaluating the success of the guidelines is mixed, undermining consensus on its effectiveness in reducing disparities (see, for example, Hofer, Blackwell & Ruback, 1999; Karle & Sager, 1991).

Researchers also focus on the impact of the *Booker* decision and the shift from mandatory sentencing guidelines to an advisory federal sentencing scheme (Hofer, 2011; Nowacki, 2015; Paternoster, 2011; Spohn, 2011; Starr, 2013; Tiede, 2009; Ulmer, Light & Kramer, 2011). The reason for this focus is clear; “if a primary goal of federal sentencing reform was a reduction of unwarranted disparity, the impact of the Booker decision on extralegal disparity is among the most important empirical questions facing sentencing policy-makers” (Hofer, 2007: 451). According to Ulmer, Light and Kramer (2011, p. 829), “the aftermath of the Booker decision provides an opportunity to examine what happens when legal decision-makers are released from relatively strong formal rational decision-making constraints, and are given more room to base their decisions on substantively rational criteria.” The United States Sentencing Commission also produced several reports evaluating changes in sentencing trends occurring subsequent to the Supreme Court decision in *Booker*. While the Commission found that the guidelines have “remained the essential starting point for all federal sentences and have continued to influence sentences significantly” (U.S. Sentencing Commission, 2012, Part A, p. 3), they also concluded that adherence to guideline standards varied notably by offense type. For example, adherence to the guidelines was diminished in fraud and child pornography

cases post-*Booker*. Moreover, divergence from the guidelines in some offense types has grown over time (U.S. Sentencing Commission, 2012).

The commission discovered that the rate of within-range sentences for most offenses decreased over time, while the rate of below range sentences increased between 2005 and 2012 (U.S. Sentencing Commission, 2012). And sentencing outcomes were found to “increasingly depend” on the federal district in which the defendant was sentenced (U.S. Sentencing Commission, 2012). Continuing differences in sentence length were also identified based on the demographic characteristics of defendants (i.e. black offenders receive longer sentences than white offenders and men receive longer sentences than women).¹⁰ The commission concluded that, post-*Booker*, “sentencing outcomes increasingly depend upon the judge to whom the case is assigned” and that appellate review “has not promoted uniformity in sentencing to the extent that the Supreme Court anticipated in *Booker*” (U.S. Sentencing Commission, 2012, p. 8).

Some independent scholars have reached different conclusions. For example, Ulmer et al. (2011) argue that sentencing disparities between districts have not worsened since the Supreme Court’s decision in *Booker*. Consequently, they conclude that vesting judges with more freedom to employ substantive rationality does not necessarily exacerbate the problem of sentencing disparities. Instead, the “norm setting function” of the guidelines has “become embedded in the organizational and legal culture of federal court communities and therefore continues to structure judicial federal sentencing” (Albonetti, 2011, p. 1153; see also Ulmer, Light & Kramer, 2011). Further, Bennett (2014, p. 490) argues that the severity of federal sentencing has not changed post-*Booker*

¹⁰ The Commission noted, however, that these results should be viewed cautiously, since not all legally relevant factors were included in the regression analyses (Department of Justice Fact Sheet, 2006; U.S. Sentencing Commission, 2012).

“because judges’ sentences are subconsciously anchored by the calculated Guidelines range.”

Despite numerous legislative efforts, sentencing disparities remain a significant problem that interferes with equitable outcomes in the criminal justice system. While there are indications that legislative measures have helped to alleviate some of the problem, sentencing disparities are still a widespread and serious concern. Also evident is the reality that judicial discretion – which will always exist to some degree – can naturally lead to disparate sentences. Given these realities, it is critical that research efforts be directed toward understanding the various ways in which judges exercise their discretion. Understanding how and where judicial discretion is being exercised improperly (i.e. irrelevant sentencing factors are influencing sentencing decisions) will not only illuminate areas of weakness that would benefit from legislative focus but can also inform well-intentioned judges of their improper (if inadvertent) use of sentencing discretion.

B. Sentencing factors

A judge’s decision to impose a particular sentence upon a convicted defendant is the product of many considerations. While certain factors (i.e. severity of the offense and the defendant’s criminal history) are legally relevant at the sentencing phase (Roberts, 1994; U.S. Sentencing Commission, 2011, p. 380; 18 U.S.C. Section 3553[a]), prior research has shown that extralegal factors (e.g. defendant’s race and gender) also unduly influence sentencing determinations (Albonetti, 1997; Everett & Wojewicz, 2002; Mustard, 2001; Rodriguez, Curry & Lee, 2006; Steffensmeir & Demuth, 2000; Ulmer, Light & Kramer, 2011). In this section, I examine the two key legally relevant

sentencing factors: seriousness of the offense and the defendant's criminal history.

Additionally, I discuss the influence of plea bargaining on sentence severity. Following a discussion of these factors, I explore a variety of extralegal factors that have been found (to varying degrees) to affect the punitiveness of sentences imposed for like offenses.

Legally Relevant Sentencing Factors¹¹

Seriousness of Offense

An important legal factor in sentencing determinations is the seriousness of the offense of which the defendant is convicted. In general, even among *mala in se* offenses, certain crime types (e.g., murder) are treated as more serious than others (e.g., theft). Illegal acts that are more serious in nature generally warrant greater punishment than minor violations of the law. While individual state penal laws and the Federal Sentencing Guidelines provide sentencing ranges based, in part, on the nature and characteristics of the particular offense committed (N.Y.S. Penal Law, Article 70; 18 U.S.C., Section 3553(a)), these guidelines vary from jurisdiction to jurisdiction (18 U.S.C., Section 3553(a); N.Y.S. Penal Law, Article 70; for example, see also, North Carolina and Pennsylvania Sentencing Guidelines). Thus, the determination of the amount of punishment that should be meted out for a particular criminal offense is not an exact

¹¹ In addition to the seriousness of the offense and the defendant's criminal history, certain aggravating and mitigating factors are also considered legally relevant sentencing criteria for some types of offenses. Much of the case law regarding aggravating and mitigating factors relates to capital cases in particular (see, for example, *Lockett v. Ohio*, 1978, pp. 604-605; Sandys, Pruss & Walsh, 2009). Different states have varying statutes with enumerated mitigating factors. For example, in Florida, mitigating circumstances include a defendant's lack of prior criminal history, the age of the defendant at the time of arrest, the fact that the victim consented to the act, situations where the defendant was an accomplice with limited minor participation in the offense and evidence that the defendant was under "extreme mental or emotional disturbance" during the commission of the offense (Fla. Statutes, 921.141(5)). Examples of aggravating circumstances include situations where the offense was committed for pecuniary gain, capital offenses that are "especially heinous, atrocious, or cruel," cases where the victim was under age twelve, and situations where defendant created "a grave risk of death to many people" (Fla. Statute 921.141).

science. Certain acts, on their face, may appear to justify more serious punishment than others (e.g., pick pocketing versus murder), yet the distinctions between other illegal actions may be much less obvious (e.g. robbery versus burglary) (see Robinson, 2008).

One of the factors widely considered in determining the seriousness of a given offense is the resulting harm to the victim (Nadler & Rose, 2003). Indeed, penal laws categorize offense severity based not only on the defendant's actions themselves but also by the harm caused by those actions – or the consequences. For example, in New York, a defendant who intends to cause physical injury to another and proceeds to cause such injury is guilty of the crime of Assault in the Third Degree. Assault in the Third Degree is a Class A Misdemeanor punishable by up to one year in jail (N.Y.S. Penal Law, Section 120.00). In contrast, Attempted Assault in the Third Degree (N.Y.S. Penal Law, Section 110/120.00), in which the defendant may have intended to harm but failed to injure another party is a Class B Misdemeanor, with a maximum sentence length three months (N.Y.S. Penal Law Section 110/120.00). The difference in potential length of incarceration between Assault in the Third Degree and Attempted Assault in the Third Degree in New York illustrates how similar actions may lead to very different legal consequences based on the resulting harm. Since harm is a relevant factor in sentencing determinations, any physical or psychological injury caused to the victim is probative information at sentencing. In fact, in all federal sentencing proceedings and in the vast majority of state sentencing proceedings, victims are permitted to give Victim Impact Statements (Cassell, 2009; 18 U.S.C. Section 3771(a)(4), 2004). This practice enables victims or their family members in federal cases to be “reasonably heard” at sentencing (Cassell, 2009). Victim Impact Statements provide sentencing judges with valuable

information regarding the harm caused to the victim and his or her family as a result of the defendant's actions. For example, a victim's statement that she has been having nightmares since her rape and has not left her home for months out of fear is helpful to a sentencing judge in assessing the degree of psychological harm experienced by the victim. This information is relevant in determining the specific sentence appropriate for the defendant (Cassell, 2009). Advocates of Victim Impact Statements argue that these statements are an integral part of the sentencing process: "It is axiomatic that just punishment cannot be meted out unless the scope and nature of the deed to be punished is before the decision maker" (ABA Guidelines for Fair Treatment of Crime Victims and Witnesses, 1983, p. 21).

Prior studies have found that both the severity of the offense and the criminal history of the defendants have the strongest influences at sentencing (Bushway & Piehl, 2001; Kramer & Ulmer, 1996). For example, in a study utilizing Pennsylvania data, Kramer & Ulmer (1996) show that offense type and the defendant's criminal history are primary influences in judicial decisions to depart from the sentencing guidelines. For the most part, there is general consensus in the population concerning perceptions of offense seriousness (Ramchand et al., 2009; Samuel & Moulds, 1986; Wolfgang, 1985); however, there is some consistent variation among groups. For example, Samuel and Moulds (1986) note that Whites view all offenses to be more severe than do other racial groups. Further, perceptions of theft are more severe in the older population, while lower income and less educated groups tend to perceive offenses involving serious physical injury to be less severe than do persons with higher income and more education (Samuel & Moulds, 1986). Finally, perceptions of offense severity are context specific; for

example, panic regarding the dangers and implications of drug use influence perceptions of the severity of drug offenses more during certain time periods than others (e.g. moral panic regarding marijuana use during certain time periods) (Ramchand et al., 2009).

Criminal History

A defendant's prior criminal record is another legally recognized and relevant factor in sentencing determinations. In fact, state and federal courts commonly utilize information from a defendant's criminal history in determining the type and length of sentence to be imposed (see, for example, 18 U.S.C. Section 3553; N.Y.S. Penal Law, Article 70). Likewise, other countries have penal codes that contain provisions for repeat offenders and prescribe more severe penalties for recidivists (Roberts, 1994).

Various rationales have been posited for the inclusion of criminal history as a legally relevant sentencing criterion. For example, the Federal Sentencing Guidelines Manual (U.S. Sentencing Commission, 2011, p. 380) provides the following rationales:

- A defendant with a record of prior criminal behavior is more culpable than a first offender and thus deserving of greater punishment.
- General deterrence of criminal conduct dictates that a clear message be sent to society that repeated criminal behavior will aggravate the need for punishment with each recurrence.
- To protect the public from further crimes of the particular defendant, the likelihood of recidivism and future criminal behavior must be considered.
- Repeated criminal behavior is an indicator of a limited likelihood of successful rehabilitation.

While the use of a defendant's prior criminal history in reaching sentencing determinations is widely accepted in legal circles, scholars disagree over whether criminal history should be relevant. A retributivist perspective argues that a defendant who has previously been convicted of a crime should receive a more severe sentence

“because the offender has already been warned and has proven himself unable or unwilling to follow society's commands” (Demleitner, 2005, p. 159). Another perspective, which engages a different logic but arrives at a similar result, posits that it is unfair to punish repeat offenders more harshly; instead, first offenders should receive a sentence discount (Demleitner, 2005, p. 159). Arguments supporting a sentence discount for first offenders often center on the idea of granting first-time offenders a second chance (Roberts, 1994).

There are many challenges involved in using criminal history at sentencing in a manner that results in fair and uniform outcomes. For example, it is sometimes difficult to ascertain how the criminal histories of immigrants, who may have criminal records from foreign countries, should be weighed (Demleitner, 2005). Often there is insufficient information available to the court on the exact nature of a foreign criminal conviction, making it difficult to evaluate how these past convictions impact sentencing determinations. Further, different countries have different burdens of proof, a factor that significantly affects the likelihood of conviction in a given case. These and other variances can frustrate the equitable use of prior convictions by sentencing courts.

The Federal Sentencing Guidelines provide a working example, however, of how criminal histories may be useful in reaching reasonably equitable sentencing determinations. In particular, the Federal Sentencing Guidelines contain an elaborate system of six criminal history categories to be used in calculating a defendant's sentence (U.S Sentencing Commission, 2011). The categories determine the number of criminal history points assigned to a convicted defendant. For example, Category 1 applies to defendants with 0-1 criminal history points while Category 6 applies to defendants with

more than 13 criminal history points. Section 4A1.1 of the Guidelines delineates how points are calculated. For example, Section 4A1.1(a) provides that 3 points be added for each prior sentence of imprisonment exceeding one year and one month.

In contrast, different state systems handle criminal histories in different ways (for an example, see N.Y.S. Penal Law, Section 70). A judge's determination of whether and how to use the prior offense(s) involves the consideration of many factors. In addition to severity of the prior offense, other factors include timing of the previous conviction¹² and the relationship between the prior and the current offense.

Plea Bargaining

Plea bargaining involves an agreement between the prosecutor and the defendant. Specifically, a prosecutor may reduce the charges brought against a defendant or propose a reduction in the type or length of sentence to be imposed in exchange for the defendant's guilty plea. Generally, a defense attorney will consider the strength of the evidence, potential sentence length, and a defendant's preference (whether he wishes to go to trial) in making plea recommendations to his client (Kramer, Wolbransky & Heilbrun, 2007). A defendant who pleads guilty relinquishes his constitutional right to a trial in exchange for the prosecutor's concessions (Edkins, 2011; Kramer, Wolbransky & Heilbrun, 2007; Walsh, 1990).

In general, then, defendants who plead guilty often receive shorter sentences than defendants who are found guilty at trial (Edkins, 2011; Kramer, Wolbransky & Heilbrun, 2007; Ulmer, Eisenstein & Johnson, 2010; Walsh, 1990). To illustrate, in 2006, over 90

¹² See, for example, N.Y.S. Penal Law, Section 70.04(1)(iv), where a prior violent felony conviction may only be used as a predicate for a second violent felony offender classification if the sentence imposed for the prior felony was within ten years before the commission of the felony for which defendant is being sentenced.

percent of offenders in state court pled guilty, and the average sentence for a defendant convicted of a violent felony at trial was five years longer than for defendants who pled guilty (Edkins, 2011). These results confirm earlier studies wherein trial defendants received longer sentences than plea defendants; for example, Walsh (1990) notes that trial defendants receive 13.55 months more prison time than do defendants who plead guilty.

Court circumstances may impact whether and the degree to which convictions following trial result in longer sentences than convictions following pleas. For example, Ulmer et al. (2010) finds that district courts that have higher caseload pressures tend to penalize defendants for exercising their constitutional right to a trial more than district courts with lower caseload pressures. They further note that the ‘penalty’ for going to trial is less in districts that have higher trial rates compared to districts that have lower trial rates (Ulmer, Eisenstein & Johnson, 2010). Undoubtedly, there are compelling practical reasons for permitting defendants to plead guilty. Avoiding a trial saves the community associated expenses, reduces court delays and removes pressure from victims to testify at trial (Bagaric & Brebner, 2002).

Extralegal Factors at Sentencing

In this section, I discuss the extralegal influence of defendant and victim demographic (race/ethnicity, gender and age) and socioeconomic characteristics at sentencing. I additionally explore how the local court, in which the defendant is sentenced, influences the severity of sentence imposed by judges.

Race/Ethnicity of Defendant

From a legal and equitable standpoint, a defendant's race should have no bearing on the sentence imposed following conviction. To the contrary, sentencing determinations should focus exclusively on legally relevant sentencing criteria, such as the severity of the offense and the defendant's criminal history. Yet, results from prior studies conclude that the race of the defendant has a reasonably strong and consistent influence on the type and severity of sentence imposed (Albonetti, 1997; Bushway & Piehl, 2001; Everett & Wojewicz, 2002; Steffensmeir & Demuth, 2000; Steffensmeier & Demuth, 2006; Ulmer, Light & Kramer, 2011).

Specifically, White defendants often receive more lenient sentences than their African American counterparts (Albonetti, 1997). For example, Steffensmeier and Demuth (2000) show that White defendants are least likely to be incarcerated in federal court and, when incarcerated, White defendants -- particularly drug defendants -- receive shorter sentences than African American and Hispanic defendants. Everett and Wojewicz (2002) likewise illustrate that Hispanic and African American defendants receive more severe sentences than White defendants in federal courts. Further, Bushway and Piehl (2001) demonstrate that African American defendants in Maryland receive sentences of incarceration that are 20 percent higher than Whites.

While the vast majority of studies address African American and Hispanic offenders, there are a handful of studies that focus on the sentencing of other ethnic groups. For example, Johnson and Betsinger (2009) conclude that Asian offenders are generally treated similarly (and at times even more leniently) than White offenders in federal district courts, and that Asian offenders receive less severe sentences than African American and Hispanic offenders. In another study, Alvarez and Bachman (1996) studied

disparities in sentence length between Caucasians and American Indians in Arizona; in that research, the authors determine that American Indians receive longer sentences than Caucasians in robbery and burglary cases but shorter sentences in homicide cases.

Some recent studies identify racial disparities in decisions regarding the type of sentence to be imposed. The U.S. Sentencing Commission (2011) and Ulmer, Light, and Kramer (2011) conclude that the decision of whether to sentence a defendant to probation or to a period of incarceration “is a source of persistent and increasing disparity” between White and African American defendants (Scott, 2011, p. 1129). More recently, Johnson and DiPietro (2012) note that male and minority defendants are less likely than White defendants to receive intermediate sanctions instead of prison sentences. Interestingly, Johnson and DiPietro also find that minority and male defendants are less likely to receive intermediate sanctions instead of probation. They suggest that these differences may exist because judges perceive male minority defendants to be “less amenable and deserving of these scarce rehabilitative programs” (Johnson & Pietro, 2012, p. 837).

Much of the research that explores the influence of race on sentencing decisions focuses on the relationship between jurisdictional characteristics and the race of the defendant. For example, Feldmeyer and Ulmer (2011) find that Hispanic defendants receive more lenient sentences than African Americans and relatively similar sentences to Whites in federal judicial districts. However, in particular, they note that Hispanics receive more lenient sentences in areas with larger Hispanic populations and the harshest sentences in districts with the smallest proportion of Hispanic residents. Accordingly, the “sociopolitical resources that accompany larger Hispanic populations may actually lead to more lenient Hispanic sentences” (Feldmeyer and Ulmer, 2011, p. 259).

There may even be a connection between the extralegal influence of race and the strength and severity of the case such that race is used as a criterion, but only in weaker cases (Devine et al., 2009; Kalven & Zeisel, 1966; Spohn & Cederblom, 1991). In their study of jury behavior, Kalven and Zeisel (1966) conclude, for example, that jurors are more likely to feel liberated to consider their own feelings and impose their own values in weaker cases. In contrast, jurors are not as susceptible to extralegal influences in more serious and stronger cases. Using Kalven and Zeisel's liberation hypothesis, Spohn and Cederblom (1991) conclude that this is because discretion is diminished (and therefore racial prejudices are muted) in more serious cases (i.e. murder, rape and robbery cases; cases where defendants have prior felony convictions; cases where the victim and defendant are strangers and cases where the defendant possesses a firearm). Likewise, judges appear to consider extralegal racial factors in weaker cases (Spohn & Cederblom, 1991).

Finally, there is evidence to suggest that the intersection between race and other demographic characteristics causes sentencing disparities to be magnified. For example, Doerner and Demuth (2010) find that Black and Hispanic defendants who are also young and male receive more severe sentences than Whites, females, and older defendants. They further show that young Black males receive the longest sentences and young Hispanic males have the highest incarceration rates (Doerner & Demuth, 2010).

In sum, prior research demonstrates that an offender's race influences the type and severity of punishment imposed at sentencing. Further, some research establishes that case severity, jurisdictional composition and other defendant demographics (i.e. gender and age) interact with race in producing sentencing disparities.

Gender of Defendant

Research shows that the defendant's gender may affect the type and length of sentence imposed by a sentencing court (Bushway & Piehl, 2001; Doerner, 2012; Freiburger & Hilinski, 2013; Mustard, 2001; Rodriguez, Curry & Lee, 2006; Steffensmeier & Demuth, 2006; Ulmer & Kramer, 1996). Generally, female offenders receive a milder type of punishment upon conviction than male offenders. This gender disparity is most pronounced in decisions of whether to incarcerate a defendant or impose a sentence of probation; females are less likely to be sentenced to prison than male offenders (Doerner, 2012; Mustard, 2001; Rodriguez, Curry & Lee, 2006). While some research concludes that females receive shorter sentences than males, the empirical results on sentence length are more mixed (Mustard, 2001; Rodriguez, Curry & Lee, 2006). Moreover, the degree of gender disparity may be dependent on the type of crime. For example, Mustard (2001) explains that gender differences are the strongest in drug trafficking and bank robbery cases and the smallest in less serious types of offenses, such as larceny, fraud, and immigration violations.

Some research focuses on potential explanations for these gender differences. For example, in a qualitative study, Ulmer and Kramer (1996) note that gender differences in decisions regarding whether or not to incarcerate a defendant are conditioned by factors including the defendant's family status and whether or not the defendant is responsible for dependent children (see also Freiburger, 2011). They conclude that "these factors lead court actors to view female defendants as less blameworthy, less dangerous, better risks for rehabilitation and thus candidates for leniency" (Ulmer & Kramer, 1996, p. 402).

Gender has also been found to interact with other extralegal factors at sentencing. For example, Steffensmeier and Demuth (2006) determine that gender has interactive effects with race/ethnicity at sentencing. They used data from large urban courts collected from the years 1990 through 1996. While their research supports prior findings that African American and Hispanic males receive harsher sentences than White males and that women receive more lenient sentences than men, they also find that sentences between White, African American, and Hispanic females are not disparate. These results suggest that any existing race/ethnicity bias in sentencing exists for male but not female defendants (Steffensmeier & Demuth, 2006).

Victim Demographic Characteristics

Some studies illustrate that the demographic characteristics of the victim influence case dispositions and sentencing outcomes. For example, Baumer and colleagues (2000) argue that both the race and gender of homicide victims influence various stages of case processing, especially when cases are decided in jury rather than bench trials. In one of the most politically influential studies describing the effects of victim characteristics on case outcomes, Baldus, Pulaski, and Woodworth (1983) show that offenders of all races are significantly more likely to receive a death sentence when the victim is White. Consequently, the sociodemographic characteristics of victims is expected to play some role in sentencing punitiveness.

Age

Age is a sentencing factor that can be legally relevant in some instances but extralegal in others. For example, age may be a legally relevant sentencing criterion in certain types of rape cases (e.g. consensual with underage victim and older defendant)

and with certain minor defendants.¹³ Further, age may be legally relevant in evaluating how a defendant's criminal history impacts sentencing decisions (Bushway & Piehl, 2007). Bushway and Piehl (2007) explain that while sentencing guidelines provide sentencing ranges based on the number of prior convictions, they do not account for the fact that older defendants have had more time to accumulate these convictions. Therefore, while two defendants of different ages may have similar criminal histories, the younger defendant may be more culpable since he or she is committing crimes at a faster rate.

The effect of offender age on sentencing decisions is somewhat inconsistent (Freiburger & Hilinski, 2013; Steffensmeier et al., 1998; Wu & Spohn, 2009). Wu and Spohn (2009) conclude that age does not impact prison sentence length and that the association between sentence length and age is very weak. Others find that age can interact with race and gender to influence sentencing determinations, particularly in cases where the defendant is young (Freiburger & Hilinski, 2013; Spohn & Holleran, 2000; Steffensmeier et al., 1998). For example, Steffensmeier et al. (1998) find that race is more influential in cases with younger defendants than in cases with older defendants. Likewise, Spohn and Holleran (2007) note that offender characteristics of race, age and gender interact to produce the most severe sentences for young, black males.

Socioeconomic Status

Prior research demonstrates that a defendant's socioeconomic status influences the length of sentence that he or she receives for his or her crimes. For example, Osborne

¹³ Age is not considered in this study. First, age is not included in the quantitative portion of this dissertation due to concerns about missing data. Specifically, data is missing in nearly one-fifth (19.4%; N=94) of the cases. Moreover, age is a sentencing factor that may be either legally relevant or extralegal, depending on the circumstances of the case. While age is not specifically explored in the qualitative section of these analyses, conversations initiated by interviewees regarding offender age were addressed and documented.

and Rappaport (1985) conducted an experiment with 193 White college students who were enrolled in undergraduate psychology classes. Subjects were given a written description of a murder case. Among two other variations (i.e. race and type of murder), summaries varied by socioeconomic status. Specifically, some summaries indicated that the defendant was a janitor while others summaries indicated that the defendant was an advertising executive. The results indicate that students advocate for longer sentences when the defendant is of lower socioeconomic status compared to higher (Osborne & Rappaport, 1985).

Prior research shows that the impact of socioeconomic status may interact with crime type in influencing sentencing decisions. For example, Wheeler, Weisberg and Bode (1982) find that high status defendants – those convicted of white-collar crimes – receive harsher sentences than low status defendants convicted of the same crimes. Drawing from interviews conducted with judges, Wheeler and colleagues (1982) explain that judges “gave evidence of strong sentiment against crimes of greed rather than need, against crimes committed by persons in positions of trust and authority” (Wheeler, Weisberg & Bode 1982, p. 657). In contrast, Benson and Walker (1988) find that high status white-collar defendants are no more likely to be incarcerated or receive longer sentences than low status defendants.

Local Court Context

Prior research shows that the local court in which a defendant is sentenced can influence the severity of sentence imposed (Johnson, 2005; Kramer & Ulmer, 2006). For example, Johnson (2005) finds that judicial adherence to sentencing guidelines vary across courts; specifically, defendants are more likely to receive sentencing guideline departures in

certain local courts than others. Defendants in larger courtrooms receive more leniency than defendants in smaller courtrooms, and courts that have more caseload pressure are more likely to employ downward departures in sentencing than courtrooms with less caseload pressure (Johnson, 2005). Johnson posits that local courtrooms differ from one another in terms of available resources and “political, social and organizational contexts” and that these differences influence sentencing outcomes because judges are hesitant to make sentencing decisions that conflict with “the cultural norms and organizational expectations of the court” (Johnson, 2005, p. 766). Kramer and Ulmer (1996) likewise note variations in sentencing departures among local courts. Specifically, there are more dispositional departures in counties that are “highly urbanized,” and defendants are somewhat less likely to receive dispositional departures in counties that have a dominant Republican electorate.

Summary

A review of the literature demonstrates that various factors influence judicial sentencing decisions. While certain factors are legally relevant and should influence sentencing determinations – such as the severity of the offense and the defendant’s criminal history – other factors should not. To date, the primary focus of the academic community in studying extralegal sentencing factors relates to the defendant’s demographic and socioeconomic characteristics or on contextual factors like locality or caseload pressures. In contrast, there has been little focus on the evidentiary characteristics of cases as a potential extralegal sentencing consideration.

C. Evidentiary Weight: A Different Brand of Extralegal Sentencing Factor

I use the term 'evidentiary weight' to refer to the type and quantity of evidence

presented during the guilt phase of a criminal case. In any given prosecution, different forms (types with varied substantive strengths) and quantities of evidence may be used to establish a defendant's guilt beyond a reasonable doubt. In this section, I define the concept of evidentiary weight more fully.

Different types of evidence may be available in criminal cases.¹⁴ For example, some of the types of evidence typically seen in criminal cases include physical and forensic evidence, witness-based (e.g. eyewitness testimony) evidentiary forms, confessions, and documentary evidence (e.g. medical records). Often, the same act may be proven, in different ways, based on the types of evidence available. For example, fingerprint analysis confirming that the defendant was at the burglarized home is an important piece of evidence in establishing a defendant's guilt of burglary. Yet this same criminal act may be established by a witness's testimony that he or she observed the defendant inside the property. Other cases may contain both fingerprint and eyewitness evidence. The fact that the evidence presented at trial includes fingerprint analysis in one case, witness testimony in a second case, and a combination of both evidentiary forms in a third case does not affect the level of culpability or the seriousness of the offense. They are merely alternative ways of establishing the same evidentiary facts. Nonetheless, different fact-finders may conclude that certain evidentiary forms – or combinations thereof – provide greater probative value than others.

¹⁴ While there is a tendency to think of evidence as independently existing in a case – waiting to be collected by the police – Mark Cooney (1994) directs us to the idea that evidence is socially produced. Cooney argues that the social structure of the actors in criminal cases is determinative of the quantity and quality of evidence that will be collected or generated by police in a given case. Specifically, he maintains that the social status of the victim, offender and other actors to the proceedings (e.g. defense attorneys, police) should influence the amount of evidence that is marshaled for and against the defendant (Cooney, 1994).

Evaluating the strength of different evidentiary forms is, in some ways, subjective. For instance, while eyewitness identification has inherent reliability concerns (Leippe, Eisenstaedt & Rauch, 2009; Wells & Quinlivan, 2008), it is still a highly influential form of evidence for jurors. In fact, for some fact finders, eyewitness testimony is the most compelling form of evidence (see Shermer, Rose & Hoffman, 2011). For others, forensic evidence is the most influential. For example, jurors in a study by Smith et al. (2011) rated DNA and fingerprint evidence as consistently more reliable than other forms of evidence, including eyewitness testimony and confessions. Some researchers argue that certain popular crime television programs, portraying criminal investigators utilizing forensic tools to solve crimes, taint juries by altering their expectations of the type of evidence required to establish a defendant's guilt at trial (Cole & Dioso-Villa, 2011). Dubbed the "CSI Effect," this argument is based on the idea that jurors who watch programs like CSI may expect or require forensic evidence to prove criminal prosecutions, and they may weigh forensic evidence more heavily than do jurors who are not exposed to these programs (Cole & Dioso-Villa, 2011; Hayes-Smith & Levett, 2011; Hughes & Magers, 2007; Kim, Barak & Shelton, 2009; Shelton, Barak & Kim, 2011). This is true despite recent literature showing that much forensic evidence fails to meet the standard of 'science' and is fundamentally flawed (Saks & Faigman, 2008; Saks & Koehler, 2005). Since the relative influence (or perceived strength) of distinct evidentiary types is not uniform, certain evidentiary packages may result in a conviction for one jury but an acquittal for another. Consequently, what constitutes a strong case varies among fact finders and is intimately tied to their perceptions of the *quality* of different evidentiary types.

The strength of the evidentiary package may be assessed, alternatively, by the number of different pieces (or the *quantity*) of evidence that prosecutors are able to marshal in the guilt phase. The combinations of various types of evidence routinely presented at trial include police witnesses, eyewitnesses, confessions¹⁵ by the defendant, and physical evidence connecting the defendant to the crime (Spohn, 2000). That is, cases that contain both forensic evidence and witness-based testimony may be more compelling than cases dependent on a single evidentiary form; cases that contain three types may be perceived as even stronger, and so on (Heinrich et. al, 2013).¹⁶

In this section, I review two broad evidentiary types that are frequently part of the evidentiary package in criminal prosecutions of violent crimes: forensic evidence and witness-based evidence. Next, I delineate why evidentiary weight should be considered an extralegal factor at sentencing. Finally, I review previous literature that suggests that some aspects of evidentiary weight – specifically, the presence of forensic evidence – influences judicial sentencing determinations.

Forensic Evidence

¹⁵ Very few forms of evidence in criminal proceedings are as influential as the uttered or written confession of a criminal defendant. As expressed by the United States Supreme Court, "a confession makes the other aspects of a trial in court superfluous, and the real trial, for all practical purposes, occurs when the confession is obtained" (*Colorado v. Connelly*, 1986). While confessions are arguably the most compelling piece of evidence a prosecutor can present, prior studies demonstrate that false confessions are not uncommon. A major cause of false confessions has been found to be certain psychological interrogation techniques employed by the police (see, for example, Gudjonsson, 2003; Inbau et al., 2001; Kassin & Kiechel, 1996; Klaver, Lee, & Rose, 2008). The effect of confessions on judicial sentencing decisions is not captured in the quantitative portion of this study for several reasons. First, in general, confessions are considered to be the most influential piece of evidence and, therefore, are expected to have the most uniform effect on judicial confidence in guilt. Second, in many cases, the existence of confessions tends to be dependent on police efforts, as opposed to case characteristics (see Cooney, 1994). Finally, data regarding confessions are not included in the public-use NIJ dataset that I employ in Chapter 4.

¹⁶ Snortum and Riva (1990) show that drunk driving convictions are influenced by the quantity and quality of evidence, and Heinrich et al. (2013) find a relationship between total number of forensic evidence items and sentence length in terrorism cases in Britain.

Forensic evidence is produced by utilizing scientific methods to investigate crimes and may be presented in many forms. For example, DNA evidence and latent prints are both forms of forensic identification evidence that "strive to achieve conclusions of individualization, the reduction of the donor pool to a single source" (Cole, 2009, p. 235).¹⁷ Other common forms of forensic identification evidence include shoe prints, firearms, ballistics, handwriting samples and bite marks. When these types of forensic identification evidence are presented in court "the goal is to link a fingerprint, writing, bite mark . . . to the one and only finger, writing, teeth in the world that made the markings" (Saks & Faigman, 2008, p.150).

While prosecutors and law enforcement generally present forensic evidence as an objective and scientific form of evidence, research shows that forensic evidence often includes speculative opinions and questionable scientific techniques (Cole, 2009; Dror et al., 2011; Dror, 2013; Edmond & Roque, 2012; Saks & Faigman, 2008). In fact, Saks and Faigman (2008) argue that some forms of forensic science (e.g. individualization sciences) fall under the category of 'non-sciences' and are characterized by the "absence of basic science origins, unsupported assumptions, exaggerated claims, lack of empirical testing and little use of scientific method" (Saks & Faigman, 2008: 168; see also Saks & Koehler, 2005). Cole (2009) suggests that forensic identification evidence is actually incapable of accomplishing its goal of definitively establishing identifications, and that claims of individualization cannot be supported by the assumption of discernible uniqueness (Cole, 2009).

¹⁷ Although DNA evidence is considered probabilistic, while fingerprints are assumed to be 'discernably unique' (Cole, 2009; Saks & Koehler, 2005).

In addition to the abovementioned weaknesses, prior studies find inconsistencies between and intra-expert regarding their analyses and results. Dror (2013) shows definitively, for example, that subjectivity plays a role in the analysis of latent fingerprints. His results reveal that there is inconsistency in latent fingerprint analysis *among* examiners and *within* the same examiners over time. Thus, even when seemingly objective statistical measures are utilized, subjectivity and bias are inherent (Dror, 2013).

Despite a significant body of research demonstrating that forensic evidence is not as objective and scientifically sound as it may first appear, the general public and the courts place a great deal of confidence in forensic science. In fact, "the trust that is laid upon the forensic sciences generally falls somewhere between uncritical faith and manufactured myth" (Saks & Faigman, 2008, p. 150; see also Edmond & Roque, 2012). Further, many erroneously believe that forensic science flawlessly applies sound basic science in its application in criminal cases (Saks & Faigman, 2008) and that an expert's opinion regarding a piece of forensic evidence (e.g. a particular fingerprint belongs to a specific defendant) is a matter of fact, rather than the opinion of the expert testifying at trial.

Further, studies conclude that judges are generally very receptive to expert opinions presented by the prosecution on behalf of the state (Edmond & Roque, 2012). Consequently, Edmond and Roque (2012) express concern that the adversarial trial system is not capable of handling scientific expert testimony in a way that ensures that defendants will receive fair trials because judges often rely on the expert's experience, without ascertaining the reliability and validity of the methods used. While judges are well versed in law, they have little training in methodology and statistics (McQuiston,

Surrett & Saks, 2009; Saks & Faigman, 2008; Wojcikiewicz, 2013). These shortfalls are exacerbated by the general inability of defense attorneys to identify or express methodological weaknesses and to access expert witnesses to testify on behalf of their clients (Edmond & Roque, 2012; Saks & Faigman, 2008). Even in jurisdictions that apply the standard for admissibility of evidence established by the Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* (1993), judges have “bent over backwards to evade the application of *Daubert* when conscientious application would lead to the exclusion of any of the nonscience forensic sciences” (Saks & Faigman, 2008, p. 163).¹⁸

In sum, while there is a significant body of literature that suggests that forensic evidence is far from an objective and factual form of scientific evidence, it is also clear that participants in the criminal justice system (e.g. jurors, judges) often treat this evidence as indisputable fact and place it on an unwarranted pedestal. This is best illustrated by the ‘CSI Effect’ in criminal case processing, wherein jurors harbor unrealistically high expectations about the occurrence and probative value of forensic evidence as a consequence of unrealistic media portrayals of technologies available to law enforcement and the probative value of forensics in criminal cases (see, for example, Cole & Dioso-Villa, 2008; Dysart, 2012). Jurors thus expect forensic evidence to be

¹⁸ In the case of *Daubert v. Merrell Dow Pharmaceuticals* (1993), the Supreme Court set new standards for determining the admissibility of scientific expert opinion in court. These factors include whether the methods used are generally accepted within the scientific community, whether they have been published and subjected to peer review, whether the methods have been or can be tested and whether the error rate is acceptable. The *Daubert* standard replaces the prior *Frye* standard in federal court. Some jurisdictions still adhere to the *Frye* standard (i.e. whether the scientific method has general acceptance in the particular field) while others follow the standard established by the Court in *Daubert*. The *Frye* standard, which is applicable in New York, treats the relevant scientific community (i.e. fingerprint examiners in cases involving fingerprint evidence) as the gatekeeper to admission of scientific evidence in the courtroom, rather than judges. This is problematic to the extent that the relevant scientific community has an incentive to reject strong criticisms of its standards and methods. *Daubert* establishes judges as the gatekeepers for the admissibility of scientific evidence and expert witnesses; this can overcome the problems associated with *Frye*, although judges do not routinely receive scientific training and thus may not be the most appropriate gatekeepers either.

presented to ‘prove’ both that the crime happened and that the defendant can be ‘objectively’ linked to it in some way. While the CSI Effect typically benefits a defendant, since a jury dissatisfied by a lack of forensic evidence may feel compelled to acquit, it can have the opposite effect when presented at trial as jurors are more likely to treat it as substantively meaningful and unfettered by subjectivity (Dysart, 2012; see also, Cole & Dioso-Villa, 2011; Hayes-Smith & Levett, 2011; Hughes & Magers, 2007; Kim, Barak & Shelton, 2009; Shelton, Barak & Kim, 2011).

Even if the data do not support a CSI Effect (Shelton, Barak & Kim, 2011), the perception that such an effect on jurors exists may alter the behavior of other players at trial – defense attorneys, prosecutors, and judges – thereby influencing the outcome of the case. For example, some judges believe that the CSI Effect is real and has impacted trials in their courtrooms (Hughes & Magers, 2007) and, in response, attorneys “are modifying their strategies to compensate for an anticipated CSI Effect” (Smith, Stinson & Patry, 2011, p. 3).

Witness-Based Evidence

Witness-based evidence – particularly eyewitnesses, police witnesses and victims – plays a central role in the prosecution of criminal cases. In many instances, the prosecutor’s case in chief is based almost exclusively on the testimony of witnesses. In fact, in certain cases, the victim is the only witness to the incident. This is especially true in crime types that are less likely to occur in the presence of others (e.g. rape) (Connolly & Gordon, 2011). Other cases contain a combination of eyewitnesses and a variety of expert witness (i.e. scientific, psychological). Still other crime types are based largely on police witnesses (e.g. undercover drug operations). Witnesses assist in the investigation

of crimes and the identification of perpetrators. Further, eyewitnesses often provide necessary probative testimony, including identifications, at trial (Loftus, 1974; Skolnick & Shaw, 2001).

In contrast to forensic evidence, which is often *perceived* as a predominantly objective and scientific form of evidence, witness-based evidence is largely subjective and is subject to more stringent credibility and reliability assessments. While credibility determinations focus on whether the witness is telling the truth, reliability determinations are focused on the accuracy of a witness' testimony (Porter & Brinke, 2009).

Determining credibility (i.e. whether or not the witness is being truthful) involves many considerations including whether the testimony sounds credible, whether the witness has a motive to lie (e.g. based on a relationship between the witness and the victim/defendant), the witness's demeanor, the witness's criminal history, and even the witness's facial expressions (Brodsky, Griffin & Cramer, 2010; Connolly & Gordon, 2011; Porter & Brinke, 2009; Wessel et al., 2006). Indeed, a fact finder's perception of a witness's credibility is sometimes affected by the emotion displayed while the witness is testifying at trial (Golding et. al, 2003; Wessel et. al, 2006). Moreover, the impact of that emotional display may be dependent on the trier of fact. For example, jurors and judges often process such emotional testimony differently. Specifically, judges, as compared to jurors, are more experienced and, therefore, more able to separate emotion from content. Therefore, degree of courtroom experience impacts the accuracy of the credibility assessments made by each group (Wessel et. al, 2006).

In some instances, the credibility of one witness at trial will be affected by the perceived credibility of an opposing witness. For example, a complaining witness who

lacks credibility may raise the credibility of an opposing witness (Connolly & Gordon, 2011). Further, the credibility of an expert witness is often evaluated by factors external to their expertise – including likability, believability, trustworthiness, and intelligence (Brodsky, Griffin & Cramer, 2010).

Credibility issues may manifest differently dependent on crime type. For example, credibility issues are often a key concern in rape cases, given biased perceptions among criminal justice actors that rape victims are not reliable witnesses (Hackett, Day & Mohr, 2008). Consequently, extralegal factors are often used by jurors in determining a rape victim's credibility, including impressions regarding the victim's appearance, the victim's social position, and the degree of emotion the victim displays on the witness stand (Hackett, Day & Mohr, 2008).

Even in situations where credibility is not an issue (i.e. where the trier of fact has determined that the witness is credible), the reliability of a witness' testimony may be questioned. For example, while an eyewitness may testify that he or she is positive that the perpetrator of the robbery is the defendant, factors external to the truthful intent of the witness may cast doubt on the accuracy of his or her testimony. Such factors may include concerns about the witness' memory, lighting conditions, the brevity in which the incident was observed, the distance from which the incident was observed, and the victim's eyesight (Loftus, 1974; Sanders, 1984; Shermer, Rose & Hoffman, 2011; Skolnick & Shaw, 2001). In these situations, even witnesses who are assumed to be truthful may be disbelieved. Indeed, a great deal of eyewitness identification testimony has been found to be inherently unreliable (Loftus, 1974; Sanders, 1984; Sharps et. al, 2007; Shermer, Rose & Hoffman, 2011). Fully 75% of convictions that were overturned

through exonerating DNA evidence were obtained, at least in part, due to eyewitness errors (Project Innocence, 2011; Shermer, Rose & Hoffman, 2011).

The circumstances surrounding lineups may also influence whether or not an eyewitness identifies the defendant. For instance, eyewitnesses may feel compelled to identify a suspect simply because the police exerted effort in staging a lineup. The witness may assume that the police assembled the lineup precisely because they believed that they had a viable suspect (Clark, 2005). Another study concluded that when witnesses learn that a co-witness made an affirmative identification from a lineup, this information impacts their lineup decision (Levett, 2013). The perceived confidence level of a witness might likewise influence credibility and reliability determinations by triers of fact (Shermer, Rose & Hoffman, 2011).

While witness-based evidence can provide strong, corroborating support in cases where the trier of fact determines that the witness is credible and reliable, the nature of witness testimony necessitates a subjective assessment of both credibility and reliability that requires individualized judgment calls on matters that are often complex and ambiguous. Yet, despite its shortfalls, eyewitness testimony is generally perceived by counsel to be a key component of the evidentiary package (Loftus, 1974; Skolnick & Shaw, 2001).

Evidentiary Weight as an Extralegal Sentencing Factor

Evidentiary weight is a critical and legally relevant factor during the guilt phase of a criminal case. Indeed, the type and quantity of evidence presented in the prosecution's case-in-chief should heavily influence trial outcomes and plea negotiations (see, for example, Cooney, 1994; Snortum & Riva, 1990; Spohn, 2000).

While evidentiary weight may vary between like cases for a host of reasons (Borg & Parker, 2001; Cooney, 1994; Roberts & Lyons, 2009), once the guilt phase of a trial is over (i.e. a guilty verdict is reached or a guilty plea is accepted by the court) and the sentencing phase commences, the specific types and quantities of evidence used to obtain the conviction lose legal relevance. The logic of viewing evidentiary weight as an extralegal sentencing characteristic grows from the well-established legal principle that, once a jury lawfully determines that the defendant is guilty of the crime charged, it is beyond the role of the sentencing judge to usurp the jury's authority by revisiting the evidence to evaluate whether the elements of the crime were proven beyond a reasonable doubt (*U.S. v. Guadin*, 1995).

The Sixth Amendment of the United States Constitution provides criminal defendants with the right to an impartial jury trial (U.S. Const. Amend. VI). The United States Supreme Court has held that the trier of fact must reach factual determinations with regard to the "elements of the crime" (*U.S. v. Guadin*, 1995, p. 510). In a jury trial, that authority rests solely with the jury (*U.S. v. Guadin*, 1995, p. 510; *Apprendi v. New Jersey*, 2000). It is well settled by the Supreme Court that the jury must establish all elements of the crime beyond a reasonable doubt. The Court's position is explained in the 2000 case of *Apprendi v. New Jersey* (2000):

At stake . . . are constitutional protections of surpassing importance: the proscription of any deprivation of liberty without "due process of law," and the guarantee that "in all criminal prosecutions, the accused shall enjoy the right to a speedy and public trial, by an impartial jury." Taken together, these rights indisputably entitle a criminal defendant to "a jury determination that [he] is guilty of every element of the crime with which he is charged, beyond a reasonable doubt." *United States v. Gaudin*, 515 U.S. 506, 510 (1995).

The term “elements of the crime” refers to the required components or parts that comprise the crime as charged that must be proven to a jury beyond a reasonable doubt. In contrast, facts categorized as “sentencing factors” (e.g. aggravating and mitigating factors) may be decided by the judge and are often associated with a lower burden of proof (*McMillan et. al. v. Pennsylvania*, 1986). For example, evidence that demonstrates the defendant’s cruelty to the victim is occasionally weighed by judges in their determinations, but only in “atypical” or “unusual” cases (U.S. Sentencing Commission, 2015, p. 6). Real offense sentencing is another option in federal courts where judges can employ upward or downward departures from the sentencing guidelines to base sentences “upon the actual conduct in which the defendant engaged regardless of the charges for which he was indicted or convicted” (U.S. Sentencing Commission, 2015, p. 6).

While real offense sentencing and evidence illustrative of unique aggravating and mitigating factors allow a judge to consider conduct outside the scope of the actual conviction charge, they do not permit a judge to arrogate the jury’s role by making factual determinations on the elements of the crime for which the defendant was convicted.¹⁹ For example, DNA and eyewitness testimony are commonly used to identify defendants as perpetrators, and identification is a necessary element of the crime. Some judges may be persuaded by the perceived strength of DNA evidence (compared to reliability concerns associated with eyewitness testimony) and, consequently, their confidence in the guilt of the accused is greater when the evidentiary package includes

¹⁹ The Supreme Court has clarified the distinction between elements of the crime and sentencing factors on several occasions. For example, in the case of *Apprendi v. New Jersey*, the Supreme Court held that the defendant’s 6th Amendment rights were violated when the judge determined facts that raised the mandatory maximum sentence beyond the maximum permitted by the jury’s verdict alone (*Apprendi v. New Jersey*, 2000). The Apprendi Court ruled that any fact that increases the maximum penalty to which a defendant may be sentenced is considered an element of the crime (*Apprendi v. New Jersey*, 2000 at 478; see also *Alleyne v. U.S.*, 2013 at 2158).

DNA. Nevertheless, the type of evidence employed by the prosecution to prove identification does not speak to the level of culpability of the offender or to the seriousness of the offense. Likewise, it does not constitute conduct outside the conviction charge nor does it capture other aggravating or mitigating factors. In this case, DNA and eyewitness testimony are merely different mechanisms for establishing guilt; neither is relevant for determining the appropriate punishment upon conviction. When judges calibrate sentences to evidentiary weight, they are essentially assuming a fact-finding role beyond their purview. Importantly, this problem will result in extralegal sentencing disparities when one defendant benefits from a judge's view that the evidentiary package is weak, while another suffers a more punitive sentence due to subjective assessments of case strength.

Yet, what little scholarly work exists on this topic indicates that guilt phase evidentiary weight plays a role in judicial sentencing decisions. After an exhaustive review of the literature, I could find only a few studies that explore the potential influence of various evidentiary forms (both forensic and witness-based) at different stages of criminal case processing, including the sentencing phase. These studies cover case processing stages ranging from arrest through sentencing; while the sentencing phase is included in these studies, little or no focus is placed on understanding why these evidentiary associations exist post-conviction.

Peterson and colleagues (1987) collected data from six U.S. jurisdictions (Chicago, Illinois; Peoria, Illinois; Kansas City, Missouri; Oakland, California; New Haven Connecticut; and Litchfield, Connecticut) for the years 1975, 1978 and 1981 wherein each criminal case in these jurisdictions was tracked from the time that the

defendant was charged through the final disposition of the case. Their research reveals that the presence of scientific evidence as part of the evidentiary package presented during the guilt phase had no effect on prosecutors' charging decisions or the likelihood of conviction. Surprisingly, the influence of forensic evidence in criminal cases is most significant at the time of *sentencing*. Specifically, the presence of a laboratory report is associated with longer sentences in four of the six cities. Cases in which a lab report had been presented at trial result in sentences that are, on average, approximately 30 months longer compared to those cases without a report (Peterson et. al, 1987).

Over 20 years later, Peterson, Sommers, Baskin, and Johnson (2010) reevaluated the effect of both witness-based and forensic science evidence on the outcomes of various types of felony cases. In this more recent study, they find that forensic evidence has some influence on case processing decisions, although "the effects of evidence vary depending upon criminal offense, variety of forensic evidence, the criminal decision level and other characteristics of the case" (Peterson et. al, 2010, p. 7). Lab-examined evidence is a significant predictor of increased sentence length in aggravated assault cases and forensic evidence linking the defendant to the crime is a predictor of increased sentence length in homicide cases. Further, cases in which a witness reports the crime to the police are associated with longer sentences, but only in aggravated assault cases (Peterson et. al, 2010). Peterson and colleagues (2010) also identify differences in the amounts of forensic evidence collected in cases based on crime type. Specifically, they conclude that forensic evidence is collected in most homicide cases but only in about one third of rape cases. In contrast, significantly less forensic evidence is collected in aggravated assault and

robbery cases (wherein fewer than 15 percent of both types of cases involved forensic lab reports).

Most recently, Peterson and colleagues (2013) again find that forensic evidence plays a role in case processing decisions. Specifically, the presence of examined forensic evidence is predictive of charges being filed, conviction at trial (but not plea agreements), and sentence length. Though not explored fully, Peterson et al. (2013) note that sentence length is shorter in situations where witnesses give reports to the police compared to cases that do not contain witness reports.

Some prior studies focus exclusively on a particular crime type in evaluating the impact of forensic evidence at different stages of the process (i.e. arrest, charging, plea/trial and sentencing). For example, in contrast to prior findings by Peterson et al. (1987), Johnson and colleagues (2012) discover that forensic evidence is not associated with increased sentence length in rape cases. In addition, research in Britain explores the use of forensic evidence in terrorism cases between 1972 and 2008 (Heinrich et al., 2013). The findings in that study suggest that there is a significant relationship between the total number of forensic pieces of evidence in a case and sentence length. Further, among types of forensic evidence, human biological evidence has the greatest evidentiary value (Heinrich et al., 2013).

A couple of other studies evaluate the influence of different types of forensic and witness evidence at various stages of criminal case processing, exclusive of the sentencing phase. For example, Baskin and Sommers (2010) find that forensic evidence is “non-determinative” in homicide cases²⁰; however, their study does not include the

²⁰ Case processing stages studied include arrests, referrals to the prosecutor, charging decisions and convictions (Baskin & Sommers, 2010).

sentencing phase of the case. More recently, Baskin and Sommers (2012) conclude that forensic evidence does not impact the likelihood of convictions in assault and robbery cases, whereas victim and witness reports are influential. However, this study does not examine the effects of evidence type and strength on the sentencing phase of the case (Baskin & Sommers, 2012).

Peterson et al.'s (1987, 2010, 2013) research speaks to the role of evidentiary factors at sentencing, yet the effect of evidentiary weight on sentencing determinations is not the focus of these studies. Moreover, differences among crime types regarding the influence of forensic and witness evidence are not explored. And while a handful of earlier studies evaluate the effect of scientifically analyzed evidence on clearance rates for offenses (Briody, 2004; Peterson, Mihajlovic, & Gilliland, 1984), almost no research focuses on the influence of various quantities and types of forensic and non-scientific data to determine their unique and combined effects on sentence length.

Even among those prior studies that suggest guilt phase evidentiary weight *may* be influencing judicial sentencing decisions, the explanations for this influence are somewhat speculative. One potential explanation is that a judge's confidence (or lack of confidence) in a defendant's guilt may be spilling over to the sentencing phase and thereby influencing judicial sentencing determinations (Gertner, 1999). This idea is illustrated in the 1989 case of *United States v. Juarez-Ortega*. In this case, co-defendants Juarez-Ortega & DeLuna were charged and tried together for two counts of distributing cocaine and one count of carrying a firearm during a drug trafficking offense. The weapons charge mandated a five year minimum mandatory consecutive sentence. While DeLuna was convicted of all three charges, Juarez-Ortega was convicted of the cocaine charges but acquitted of the weapons charge. Nonetheless, the judge

imposed identical sentences upon both defendants. In justifying the sentence, the judge stated (*U.S. v. Juarez-Ortega*, 1989, p. 748-749):

Well, I'll tell you something. I have been disappointed in jury verdicts before but that's one of the most important ones, because what it did, it set up a disparity in result between the two defendants . . . This firearm was used. They (the jury) had to absolutely disregard the testimony of the government agent for no reason – no reason.

Thus, the judge's disagreement with the jury's verdict influenced the sentence imposed and thereby led to sentencing disparities (Gertner, 1999; Lear, 1993), such that both were subject to the same sentence yet convicted on varied counts. The outcome of *U.S. v. Juarez-Ortega* demonstrates that the level of confidence a judge has in the correctness of a jury's guilty verdict may influence the severity of the sentence imposed. Specifically, evidentiary weight (whether strong or weak), presented during the guilt phase of a criminal proceeding, may play a critical role in sentencing determinations.

Disentangling the extralegal role of evidentiary weight from legally relevant factors in sentencing determinations is important for several reasons. First, the right to a jury trial – in most felony and some misdemeanor criminal prosecutions – is a fundamental right, secured by the 6th Amendment. A defendant convicted at trial on the basis of weaker evidence is no less culpable in the eyes of the law than a defendant convicted by a stronger evidentiary package. A judge who is influenced by evidentiary weight at sentencing is, in essence, revisiting the trier of fact's determination of guilt. The extralegal use of evidentiary weight at sentencing may therefore lead to disparate results that are arguably unjust. Paradoxically, in convictions on weak evidentiary packages, judicial discretion may be erroneously used as a 'corrective' measure at sentencing to right perceived injustices that occurred at the trial level. Yet this 'corrective' measure is

substantively unjust when it fails to move judges to simply overturn guilty verdicts that are based on negligible or problematic evidence.

The Case of Pleas

The role of evidentiary weight at sentencing should vary in trial and plea cases. In trial cases, the sentencing phase is a distinct stage of case processing in which evidentiary weight should have little or no influence. A guilty verdict has already been rendered and the question of whether or not the evidentiary package warrants a conviction has been answered; a particular sentencing judge's agreement or disagreement with the conviction, based on perceptions of evidentiary strength, is not a valid sentencing factor (see *U.S. v. Gaudin*, 1995). Instead, legally relevant factors (e.g. seriousness of offense, criminal history) should be the key determinants.

By contrast, the determination of the conviction charge and the type and length of sentence the defendant will serve is generally evaluated concurrently in plea cases. In those instances, evidentiary strength is relevant to determine the charge *and* sentence of an acceptable plea, since court actors are influenced by the probability of conviction at trial in determining the appropriate plea and sentence to offer or accept (Bushway, Redlich & Norris, 2014). Specifically, the extent of the “bargain” offered to defendants may be influenced by the quantity and quality of evidence available to prosecutors, wherein cases with weaker evidence can generate larger sentence and/or charge discounts because the perceived likelihood of conviction at trial – referred to as the ‘shadow of the trial’ model – is lower (Kramer & Ulmer, 2002; see also Bibas, 2004).²¹ Consequently,

²¹ Recently, Bushway and colleagues (2014) provided an empirical test of the shadow of the trial model, based on an online survey of judges, prosecutors and defense attorneys. They tested the model, using a hypothetical case scenario (in which types of evidence and criminal histories varied), by assessing the

sentences in plea cases are generally reflective of prosecutorial discretion and are intertwined with perceptions of case strength, whereas sentencing post-trial is predominantly a reflection of judicial sentencing discretion (Johnson, 2003) and occur after guilt determinations have been reached.

relationship between the expected sentence at trial (calculated by multiplying the probability of conviction with the expected sentence at trial) and the acceptable plea. They found that prosecutors and defense attorneys generally acted in a manner consistent with the shadow of the trial model, while judges did not. Instead, the judges' behavior was "consistent with a model that offers fixed discounts to people who plead relative to what they could expect to get at trial" (Bushway, Redlich & Norris, 2014, p. 750). In general, they found that case evidence was "particularly effective at moving probability of conviction and not very important for the sentence at trial" (Bushway, Redlich & Norris, p. 741).

CHAPTER 3: METHODOLOGY

This dissertation employs a mixed methods approach. The combination of quantitative and qualitative methods to address questions regarding the factors that influence judicial sentencing decisions is ideal for several reasons. In order to examine the associations between various quantities and forms (i.e. forensic, witness based) of evidence and the severity of sentences imposed upon convicted defendants, a quantitative approach is employed in the first analytic chapter of this dissertation (Chapter 4). This allows for an investigation of patterns among hundreds of cases that would not be possible using qualitative methods alone. Findings from these quantitative analyses, that focus on the existence and degree of this association, is a helpful first step in assessing the presence and scope of the problem; that is, whether evidentiary quantity and type act as extralegal sentencing factors that contribute to sentencing disparities.

The results of these quantitative analyses suggest that judicial discretion at sentencing is heavily influenced by evidentiary quantity and type. Yet, inferences regarding how and why these factors operate to produce sentencing disparities are largely speculative. Said differently, the pattern of quantitative results is suggestive of the process linking evidentiary weight to disparities in sentencing, but of necessity is tentative. For example, one could argue that a finding in which the introduction of forensic evidence at trial is associated with longer sentences suggests that judges possess greater confidence in a defendant's guilt when the direct evidence is viewed as scientific and thus more reliable. While logical, this explanation involves a speculative leap that goes beyond the scope of the results. Therefore, qualitative research methods, in which in-depth interviews are conducted with state court judges empowered to try and sentence

defendants on violent felony offenses, are utilized to illuminate the explicit *decision-making processes* and *rationales* that help to explain these associations. Since judicial thought processes at the sentencing stage are required to confirm these speculative explanations, first hand interviews with judges provide unique and invaluable insight that is only available through direct dialogue with sentencing decision makers. In the sections below, I describe both the quantitative and qualitative components employed in this study.

A. Quantitative Analyses

The quantitative analytic chapter of this dissertation explores whether and to what degree the quantity and type of evidence used to establish guilt during the guilt phase of a criminal case influences sentence determinations post-conviction. Analyses estimate a measure of the total quantity of physical evidence in a case on the length of custodial sentences imposed upon convicted defendants, in addition to distinguishing by evidence type. In accordance with prior studies (Peterson et al., 1987, 2010 & 2013)²², it is expected that violent felony cases in which the direct evidence includes a forensic laboratory report and those cases with one or more eyewitnesses will lead to longer prison sentences for convicted defendants. Despite some reliability concerns, both types of evidence can be highly influential for the prosecution's case (Loftus, 1974; Shermer, Rose and Hoffman, 2011; Skolnick and Shaw, 2001). In addition to re-examining the influence of forensic and eyewitness evidence on sentence length in violent felony cases, I add to these expectations drawn from the extant literature a set of formal hypotheses

²² Peterson et al.'s 2010 and 2013 studies are based on the NIJ dataset used in this study. However, Peterson et al.'s studies do not test for the quantity of physical evidence in a case or study sentence length for violent crimes (i.e. aggravated assaults, rapes, robberies and homicides) exclusively in trial cases.

around how the quantity of physical evidence in a case operates post-conviction, and how the quantity and type of evidence operate post-conviction in trial cases specifically.

Therefore, the first hypothesis tested in the quantitative analyses is as follows:

H1: Defendants convicted in violent felony cases will receive longer prison sentences when the cases contain more pieces of physical evidence.

Disentangling the role of evidentiary weight at sentencing post-trial from sentencing in plea cases is critical in studying the influence of evidentiary strength on judicial sentencing decisions. As opposed to plea cases, in which guilt and sentencing decisions are often made concurrently, post-trial sentencing is a distinct stage of case processing in which the question of guilt has already been resolved. Moreover, as opposed to plea cases, in which the prosecutor is afforded considerable discretion, judges have the greatest sentencing discretion in the post-trial arena (Johnson, 2003). Trial cases provide judges with considerable opportunity to view the evidence and evaluate its strength. In many jurisdictions, wide sentencing ranges grant judges significant amounts of discretion for many offenses (see, for example, N.Y.S. Penal Law, Section 70). Consequently, any effect of evidentiary weight at sentencing post jury conviction may generate disparate and unjust results. The following set of hypotheses are tested in post-conviction trial cases:

Defendants convicted in violent felony cases will receive longer prison sentences when:

H2: The guilt-phase evidence includes a forensic laboratory report.

H3: The guilt-phase evidence includes one or more eyewitnesses.

H4: The guilt-phase evidence contains more pieces of physical evidence.

The quantitative analyses chapter utilizes data from a study entitled “*Impact of Forensic Evidence on the Criminal Justice Process in Five Sites in the United States, 2003-2006.*”²³ The data consist of a random sample of 4,205 criminal cases/reported crime incidents from Los Angeles County, CA, Indianapolis, IN, Evansville, IN, Fort Wayne, IN, and South Bend, IN between the years of 2003 and 2006, stratified by jurisdiction and crime type. The data are culled from prosecutor case files, crime laboratory reports, and police incident and investigation reports. Since this study explores the influence of evidentiary weight in violent felony cases at the post-conviction stage, a subset of cases, in which defendants have been *convicted* of *homicide, rape, robbery, or aggravated assault*, is utilized from the larger dataset. Of the 516 cases involving convictions on violent felonies, 23 (4.46%) were missing data on key variables. In the remaining 493 valid cases, 9 (1.83%) defendants received probationary sentences. Sentences of probation in violent felony cases are both qualitatively distinct and exceedingly rare, and thus are excluded from the analyses (King and Zeng, 2001). The remaining 484 defendants received an in-custody sentence ranging in length from 1 to 2,760 months (or 230 years).²⁴

Ordinary Least Squares regression models are employed to predict sentence length, which is measured as an overdispersed count variable. In practical terms, there is little qualitative difference (from a defendant’s perspective) between sentences of 40 years and significantly longer sentences (i.e. sentences of over 100 years). However, the dependent variable is not truncated to 40 years in these latter cases because the length of

²³ Available at www.icpsr.umich.edu/icpsr/ICPSR/studies/29203/documentation

²⁴ The dataset does not code multiple offenses within the same incident; therefore crime type distinguishes the most serious offense (or top charge) on which the defendant is convicted (D. Baskin, *personal communication*, May 31, 2015).

the imposed sentence measures the theoretical severity of judicial sentencing determinations. In effect, extreme sentences are suggestive of greater judicial punitiveness at sentencing and provide highly probative information regarding judicial mindset. The dependent variable represents a count of the number of months that the convicted defendant was sentenced to a period of incarceration, or *sentence length*. Since the sentence length variable is highly skewed, I transformed it by taking its natural log. OLS regression is appropriate because it provides an ideal method to estimate the relative strength of the effects of the various legally relevant and extralegal factors on the length of sentences imposed upon convicted defendants.

In order to test the role of evidentiary weight at sentencing, three evidentiary variables are included in the models: quantity of physical evidence (a count variable measuring the number of pieces of physical evidence in a case), the presence of a forensic lab report, and the presence of eyewitnesses.²⁵ Additionally, three variables reflect legally relevant sentencing characteristics; these include offense seriousness, the defendants' criminal history, and mode of conviction (plea versus trial). Finally, traditional extralegal sentencing factors are controlled in the model. These include the defendant's race and gender, the victim's race and gender, and the site in which the sentence was imposed.

In total, the quantitative analytic chapter will report the results of two OLS regression models exploring the influence of evidentiary weight on sentence length: The first model utilizes the full sub-sample of criminal cases to test the influence of the quantity of physical evidence (to test H1) as well as the relative influence of eyewitness evidence and forensic evidence on judicial determinations of sentence length (to re-

²⁵ A detailed description of the operationalization of the independent variables is provided in Chapter 4.

examine prior study findings); The second model predicts the influence of evidentiary type (i.e. forensic v. witness) and quantity on sentence length in post-conviction *trial* cases alone (to test H2, H3 & H4).²⁶

This dataset is unique in that it includes numerous and varied evidentiary variables, defendant and victim characteristic variables, and case processing (i.e. stages of processing) variables. It therefore provides important data that speaks to questions of the role of evidentiary weight in sentencing determinations currently unavailable elsewhere. However, the quantitative methodology employed in this study has several limitations. First, the measure of witness-based evidence is blunt (i.e. data includes eyewitnesses but does not include defense witnesses, confessions, expert witnesses or distinguish between police and civilian witnesses) and should be broadened to incorporate expert witnesses, police witnesses, character witnesses, alibi witnesses and the like, to capture a more complete picture of witness-based evidence in the case.

In addition, while the data is stratified by crime type, there is no information regarding whether pleas were taken to the top charge or a lesser offense. This information is important in accurately assessing the strength of any observed associations at sentencing. Specifically, different levels of a given offense often have distinct indeterminate sentencing ranges. Knowledge of the sentencing range for a specific offense is instrumental in evaluating the influential degree of evidentiary weight.

²⁶ Results of a plea only model are reported in Table 4.5. Since guilt and sentencing decisions in plea cases often occur concurrently, and sentence length in pleas are primarily a reflection of prosecutorial discretion, they are not nearly as probative in studying the influence of evidentiary weight on *judicial* sentencing discretion as trial cases. As Table 4.5 shows, the only statistically significant evidentiary variable is lab-examined evidence. Regression results reveal that sentence length in violent felony plea cases is 53.6% longer when the evidentiary package includes a forensic lab report.

Further, more specific information on the plea process is necessary to ascertain the extent to which charge bargaining, sentence bargaining, or both are incorporated into the conviction charges in different jurisdictions; these kinds of bargaining may have substantive effects on empirical analyses of back-end case processing when, for example, offense seriousness is systematically underestimated in the context of charge bargaining.

Finally, the analyses are limited to convictions that resulted in periods of incarceration and does not include cases involving probationary sentences. The felonies examined in this study are serious and violent offenses, against the person of another, resulting in custodial sentences in more than 98 percent of the cases. Therefore, the prevalence of probationary sentences is very small; nonetheless, examining the role of evidentiary weight in predicting sentences of probation versus incarceration represents an important area for future research.

While the quantitative portion of this study tests the presence and scope of the influence of evidentiary weight during sentencing, questions of how and why evidentiary quantity and type influence sentence length is open to interpretation. A potential explanation is that the impact of evidentiary weight is operating through its influence on judicial confidence in guilt. For example, forensic evidence may increase a judge's confidence in a defendant's guilt and thereby encourage the application of lengthier prison sentences. To explore these questions further, this dissertation explores and contextualizes the quantitative findings based on qualitative interviews of New York State Supreme and County Court judges.

B. Qualitative Analyses

The qualitative component of this study explores how and why evidentiary weight and type influence judicial sentencing decisions. To that end, in-depth interviews were conducted with 41 State Court sentencing judges. The strategy of purposefully sampling judges involved in the sentencing process can “provide information that can’t be gotten as well from other choices” (Maxwell, 2005, p. 88). For purposes of this study, judges are ideal interview subjects since they are experts in sentencing, exposed to numerous and varying types of cases, and uniquely able to illuminate the thought processes and rationales by which they arrive at their sentencing decisions.

Sample

I recruited judges presiding in Supreme or County Court, Criminal Term, in New York State. Among other responsibilities, these judges preside over felony jury trials, determine guilt in bench trials, accept guilty pleas, conduct hearings, and sentence defendants appearing in their courtrooms. They are exposed to distinct evidentiary forms (i.e. forensic, witness-based) and preside over cases with varying levels of evidentiary strength. In order to qualify for this study, judges were required to have felony trial experience as a Supreme or County Court Criminal Term judge in New York State.

In total, I interviewed 41 judges from 18 different counties throughout New York State. As Table 3.1 shows, 22 judges presided in urban counties, 8 judges presided in suburban counties, and the remaining 11 judges presided in rural counties.²⁷ As Table 3.1 shows, judges varied in terms of their legal backgrounds in that 17 judges were former prosecutors, 10 had defense backgrounds, 12 judges had both prosecutorial and defense

²⁷ Some of the counties are located in extremely rural areas that only have one or two criminal term county judges. In order to avoid identifying any particular judge as a participant, the names of the upstate counties are not listed.

experience, and 2 were law secretaries prior to taking the bench. Further, 26 judges were elected to the bench and 15 were appointed. Experience on the bench ranged from less than 2 years to over 30 years. Despite their differences, all judges shared a critical similarity in that they presided in New York State and were mandated to comply with the New York Penal Code, New York Criminal Procedure Law, and judicial ethical guidelines.

[INSERT TABLE 3.1]

Accessing Judges

I established initial contact with judges using one of four methods: personal contacts, cold-calling, a hybrid method incorporating professional introduction letters and cold-calls, and referrals from other interviewees (i.e. snowball sampling). As a former prosecutor, I was able to access a few close contacts and some distant connections to the judiciary. My contacts were concentrated predominantly in one urban county and, therefore, were insufficient to obtain a diverse sample of state judges or the required number of participants. Thus, multiple recruitment methods were necessary to achieve my objective. In total, I recruited twenty judges through personal contacts and referrals, five judges through the hybrid method, and sixteen judges via cold-calling methods alone.

I began the process in June of 2014 by approaching several personal contacts because I believed these contacts would provide the easiest access to a few judges and, consequently, would increase my legitimacy by establishing a successful track record of accessing participants (Goldman & Swayze, 2012; Lofland, Snow, Anderson & Lofland,

2006). I was successful in scheduling several interviews within a couple of weeks; others took several months to complete due to the judges' summer vacation schedules and delayed responses by contacts to requests for assistance.

Given the slow pace in setting interviews through personal contacts, I decided to utilize all four recruitment strategies simultaneously. First, I began cold-calling after the first few weeks of recruiting efforts. Cold-calling involved many steps including identifying judges who fit the criteria, obtaining chambers contact information, preparing a presentation for both judges and their staff and building relationships with numerous gatekeepers.

I accessed the names of eligible judges largely through online research. Specifically, publicly available websites list all trial court judges, contain contact information, and list some professional and educational background information. While additional research was necessary to ascertain applicable background details in certain instances (e.g. trial experience), this information was obtained using a combination of online research (i.e. court websites and internet search engines) and direct communication with courthouse personnel.

Initial contact with five judges was accomplished utilizing a hybrid method. Essentially, one of the judges who I interviewed sent an email correspondence to all judges in that county introducing my study and forwarding my interview request. I then contacted all of the judges in that county, via phone, referencing the email. While many of the judges did not remember seeing the correspondence or had quickly read and deleted the email, the reference to that communication lent credibility to my study and was instrumental in securing five interviews.

Whether the interview was obtained via a personal contact, cold-call or the hybrid method, I made it a practice to always ask for a referral at the end of each interview. Judges varied in their reactions. Over half of the judges interviewed provided referrals and four judges personally contacted prospective interviewees on my behalf. Five judges articulated hesitancy to refer because they did not want their colleagues to feel pressured to comply. In total, the recruitment and interviewing process took 7 months between June of 2014 and January of 2015.

Presentation of Self, Study and Ethical Concerns

During the first moments of communication with judges and their staff, I established my credentials by informing them of my university affiliation, former position as a prosecutor, and current status as a PhD candidate collecting data for her dissertation. I explained that I needed to interview judges in order to complete the data collection process and that I was hoping they could help me by participating in my study. I then presented a brief description of the study. Specifically, I explained that I was hoping to gain a better understanding of the thought processes and opinions of judges with regard to sentencing practices by conducting confidential interviews with state court judges who handle felony cases and possess trial experience. I further explained that I was familiar with judicial ethical guidelines and that my questions would remain within proscribed parameters by avoiding inquiries into specific cases. I delineated some of the topics I was exploring, offered to provide a copy of the applicable judicial ethics section²⁸ and noted that the study was approved by my university's Institutional Review Board

²⁸ See New York Advisory Committee on Judicial Ethics, Opinion 11-138. In one of the earliest interviews, a judge handed me a copy of Opinion 11-138 and suggested that I show it to any judge who states that he or she is not permitted to speak with me.

which oversees human subjects research. I then indicated the number of interviews that I had completed or scheduled to date.

A critical step in accessing judges involved predicting and pre-emptively addressing concerns regarding the risks of participation. I emphasized that the study was confidential and any published results would not include specific attributions. I further assured judges that they could elect to skip any question and terminate the interview at any time. I described the procedures utilized to safeguard the data (e.g. the storage of data on password protected computers, redactions of personal identifiers, destruction of audio-recordings after a specified time period) and provided judges with an IRB approved consent form highlighting and restating these protections. A copy of the consent form provided to judges is included in Appendix 3.1.

Apart from confidentiality, I explained that the interview was designed to last approximately one hour and that I would meet the judge in his or her chambers at his or her convenience, thereby minimizing the time required for participation. Judges varied in their time constraints. Some judges met me in the morning or late afternoon, had relatively clear calendars, and sat with me for hours virtually uninterrupted. Other judges squeezed me into their calendars during the court lunch break, fielded calls from the courtroom during our meeting, and were under pressure to conclude the interview by a certain time.²⁹ Most judges were eager to vent their frustrations and to describe their experiences. This led to prolonged interviews in numerous situations. In the end, the average interview lasted approximately 90 minutes.

²⁹ One rural judge, who performed many functions in addition to criminal cases, took a break from the interview to perform a marriage ceremony.

Interview Strategy

A range of interviewing strategies and techniques were utilized to comprehensively study the sentencing process as understood by judges themselves, as well as to explore the influences on decision making and relevance of criteria influencing sentencing determinations. In this section, I describe the interviewing strategy, including the general interview guide approach (Patton, 2002) and the use of a series of hypothetical case scenarios or vignettes.

Interview Guide Approach

An interview guide approach was utilized in these interviews. This approach “lists the questions or issues that are to be explored in the course of an interview . . . to ensure that the same basic lines of inquiry are pursued with each person interviewed” (Patton, 2002, p. 343). Therefore, the topics and issues that each interview will cover were specified prior to the interview in “outline form” (Patton, 2002). This approach allows the interviewees to “speak in their own terms” while still ensuring that all necessary topics will be covered in a “guided conversation” (Lofland et al., 2006). Additionally, an interview guide with written questions was prepared as a loose guide to help navigate the interviews, where practical. A copy of the interview guide is included in Appendix 3.2.

Several topics were addressed in all interviews. These subject matters include:

1. Sentencing factors
2. Judicial discretion
3. Perceptions of different evidentiary forms
4. Attitudes regarding jury competency

Each of these areas is discussed in the sections below.

Sentencing Factors

Using general, open-ended questions, judges were asked about different factors they consider in rendering a sentence (see section III in Appendix 3.2). Judges were asked to explain their entire thought process, including considerations that may seem minor or insignificant. Where practical, I avoided legal language in posing my questions. The goal of exploring these factors in an open-ended manner using language devoid of legalities was to focus judges on the factors that they feel are important and not factors that are legally “supposed” to be important. Veering the conversation toward the interviewee’s personal opinions and feelings was instrumental in creating a dialogue whereby judges were more open and willing to explain their decision making process and rationales. When necessary, I used specific probes in order to ensure that certain areas were covered.

Judicial Discretion/Sentencing Disparities

A segment of the interview guide was designed to explore judges’ attitudes and beliefs regarding judicial discretion (see Appendix 3.2). For example, do judges believe that they have too much or too little discretion at sentencing? How does judicial discretion during sentencing impact the system as a whole? Does the interviewee believe that judicial discretion can and/or should be used to right injustices that occurred at the trial level?

Judges were also asked questions regarding their feelings about sentencing guidelines. Should judges be mandated to follow rigid guidelines? Should all guidelines be advisory in nature? Are mandatory minimum sentences beneficial or harmful? Judges

were additionally probed on their perceptions regarding sentencing disparities. What factors come into play in creating these disparities? What can be done to reduce sentencing disparities?

As with other topics, questions regarding judicial discretion and sentencing disparities focused on how judges feel, not how judges act. As opposed to action questions, feeling questions generally do not require justification whereas action questions may provoke a defensive response that can limit the amount and quality of data generated by the interview (Patton, 2002).

Perceptions of Various Evidentiary Forms

Judges were questioned regarding their opinions on the strength and reliability of different evidentiary forms (see Appendix 3.2). Specifically, judicial opinions regarding different forms of forensic evidence (DNA, fingerprints, chemists, ballistic evidence) and their relative strength to one another were explored. Judges were also questioned about their views of witness testimony, including their perceptions of eyewitness reliability and credibility, character witnesses, alibi witnesses and expert witnesses. Judges were additionally asked to compare the strengths and weaknesses of different evidentiary forms.

Attitudes Regarding Juries/Perceptions of Guilt-Phase Verdicts

A judge's confidence or lack of confidence in a jury's ability to render a just verdict can influence sentencing decisions (See *U.S. v. Juarez-Ortega*, 1989). While the sentencing phase occurs after the jury determines guilt (in a jury trial), judges who lack confidence in the jury's decision may, inadvertently or purposefully, insert their own

impressions regarding degree and certainty of guilt into their sentencing determinations. In contrast, judges who believe that juries are generally well equipped to handle their fact-finding role may be less inclined to revisit evidentiary issues that have already been determined by the jury and are legally irrelevant at sentencing.

Various issues relating to judicial attitudes regarding juries were explored in Section V of the Interview Guide. More specifically, I posed questions along the following general lines of inquiry: Do you feel that most juries perform competently in evaluating the evidence and rendering a verdict? Do you feel that jury trials result in injustices? Have you ever felt powerless to prevent injustices during the guilt phase of a case? Have you ever felt the desire to mitigate an unjust guilty verdict after a determination of guilt? In what way? And have you ever overturned a jury's guilty verdict or felt uncertain of their own verdict in a bench trial?

Informal Conversational Interview Approach

At times, questions posed during the interview prompted judges to reminisce about their own judicial experiences of like cases. In those situations, it was advantageous to “go with the flow” and be responsive to concepts introduced by different judges. The informal conversational interview approach, as described by Patton (2002, p. 342), “offers maximum flexibility to pursue information in whatever direction appears to be appropriate.” While this approach allowed for the interview to be tailored toward the needs of each judge, it is less systematic than other portions of the interview and resulted in more difficult data analysis (Patton, 2002).

Vignettes

Vignettes are used in social science research to study a wide range of social issues (see, for example, Barter & Renold, 2000; Clancy et al., 1979;³⁰ Hills & Thomson, 1999; Hughes, 1998; Phillips, 2008). A vignette is a story that provides a selective “snapshot” of a specific situation. After being provided with the vignette, study participants are generally asked questions regarding how they would respond to the given situation and their opinions are sought about the circumstances presented in the vignette (Hughes, 1998). Prior research has shown that utilizing vignettes “has the ability to capture how meanings, beliefs, judgments and actions are situationally positioned” (Barter & Renold, 2000, p. 308).

Vignettes were utilized in 22 of the 41 interviews.³¹ With the exception of Interview #1, the vignettes were presented at the end of the interview.³² A series of four written vignettes were shown to participants. The vignettes describe details of a case in which the defendant was convicted of robbery or assault, as well as other details. After reading the vignette, judges were asked to impose a sentence upon the defendant based on the details described in the vignette. In this section, I describe the various factual components specifically included and excluded in each vignette and explain the rationale

³⁰ This study explored consensus/dissensus among federal judges in 16 hypothetical cases.

³¹ The vignettes were omitted in 19 interviews for a variety of reasons. Some judges stated that they did not want to participate in the vignette portion of the interview. Specifically, a few judges stated that they felt uncomfortable rendering a sentence based on the information contained in the vignette. Four judges stated that they were either not sure or not allowed to answer hypothetical questions (despite being provided with a copy of the judicial ethical guidelines regarding interviews with graduate students). Several others politely declined without providing an explanation. Timing constraints prevented the use of the vignettes in eight interviews.

³² In Interview #1, the vignettes were posed to the judge immediately after we finished discussing the judge’s judicial and legal background. At the end of the interview, I asked the judge if he had any suggestions for me regarding the interview process. He stated that it was “off putting” to be asked to render a sentence on a hypothetical “right out of the box.” He suggested that I move the hypotheticals to the end of the interview so that judges have the opportunity to “get to trust me” and that I don’t “scare people away immediately.” Based on this judge’s suggestion, I moved the vignettes to the end for all future interviews.

for their inclusion or exclusion. I also discuss the benefits and limitations of using vignettes as a research tool for this study.

Inclusions

Each vignette (all vignettes are included in Appendix 3.3) describes a case in which the defendant is convicted of either Robbery in the First Degree³³ or Assault in the First Degree³⁴ under New York State law. These crime types were selected for three reasons. First, under New York State law, both Robbery in the First Degree and Assault in the First Degree are violent B felonies, a crime category that permits a judge to sentence the defendant to a determinate sentence of 5 to 25 years (see N.Y.S. Penal Law, Sections 70, 160.15 and 120.10). In order to maximize potential variations in sentence, it is helpful to use a crime type in which a judge has a wide range of possible sentences that he or she can impose.

Second, robbery and assault are crime types capable of being established by different evidentiary packages and, as such, hypothetical fact patterns may be created using various quantities and forms of evidence. For example, to establish guilt of Robbery in the First Degree under New York Penal Law Section 160.10 (3), the prosecution must prove that the defendant “used or threatened the immediate use of a

³³ Under New York State Penal law, Section 160.15, “a person is guilty of robbery in the first degree when he forcibly steals property and when, in the course of the commission of the crime or of immediate flight therefrom, he or another participant in the crime: 1. Causes serious physical injury to any person who is not a participant in the crime or 2. Is armed with a deadly weapon or 3. Uses or threatens the use of a dangerous instrument or 4. Displays what appears to be a pistol, revolver, rifle shotgun, machine gun or other firearm. See, also, N.Y. Penal law Section 120.10 (Assault in the first Degree).

³⁴ Under New York State Penal law, Section 120.10, “a person is guilty of assault in the first degree when: 1. With intent to cause serious physical injury to another person, he causes such injury to such person or to a third person by means of a deadly weapon or a dangerous instrument. Only subdivision 1 will be utilized in the vignettes.

dangerous instrument” during the commission of the robbery.³⁵ The “dangerous instrument” element may be proven in different ways. In some cases, a dangerous instrument is recovered. In other cases, the prosecutor can successfully establish this crime element through eyewitness testimony alone. Some cases may contain both witness and forensic evidence. Moreover, the quality of the evidence may vary from case to case. Given the range in the quantity and type of evidentiary packages available, robbery and assault are strong choices to use in the vignettes. Finally, both robbery and assault cases are represented in the quantitative analyses.

The vignettes describe elements of the crime, including details of weapons used and items stolen. In addition, a description of the harm caused to the victim and the defendant’s criminal history are provided. In all scenarios, interviewees were informed that convictions were obtained via jury verdict. The vignettes highlight the nature and strength of the evidence presented at trial.³⁶ Specifically, the description includes witness testimony, as well as physical, forensic or documentary evidence. Judges were further informed that the defendant did not confess to the police.

The facts relating to the type and strength of evidence presented in the case vary among the four vignettes. Scenario A presents a case that contains neither forensic evidence nor eyewitnesses other than the victim. Scenario B presents a case with both forensic evidence and eyewitness testimony. Scenario C presents a case that contains eyewitness testimony but no forensic evidence. Finally, Scenario D presents a case that contains forensic evidence but no eyewitness testimony.

³⁵ For a definition of dangerous instrument under N.Y. Penal Law, see section 10, subdivision 13.

³⁶ All four vignettes describe cases in which the conviction was obtained following a jury trial. Since a judge’s confidence in guilt (and in the jury’s ability to reach just outcomes) is a focal point of this study, trial cases may provide more probative information than cases where the defendant was convicted as a result of a guilty plea.

By design, the legally relevant sentencing factors (criminal history, felony level and degree, harm to the victim and disposition type) are substantially similar in all four cases. Crime type is the one exception. Specifically, two of the vignettes describe robbery scenarios and two describe assault scenarios. A critical goal in designing the vignettes was to minimize variations among the legally relevant sentencing factors so that results highlight whether and how the quantity and type of evidence in the case influences judicial sentencing determinations. However, using four similar fact patterns with identical offenses and only varying evidentiary quantities and forms may significantly impair the efficacy of the interviews by priming the judges to the study's purpose. Therefore, two different crime types were selected.

Given the design of this study, the drawbacks of using both assault and robbery for crime types are minimized. First, Assault and Robbery in the First Degree are both Class B violent felonies in New York with the same sentencing ranges. More significantly, judges were questioned about the factors considered in their analyses. Each vignette was addressed independently and contrasted with the other three vignettes. This method allowed judges to express their perceptions and opinions about each vignette and to explain the differences they perceive among the vignettes. It is not merely the sentence rendered that is important. Indeed, it is the dialogue that the vignettes generated regarding judicial perceptions and opinions that is even more insightful. Further, two vignettes – with extremely similar legally relevant factors – were given to judges for each crime type, thus creating a basis for comparison within crime type.³⁷

³⁷ Homicide and rape cases were not chosen as the crime type for the vignettes for several reasons. First, the vignettes are designed to vary as little as possible in all legally relevant and extralegal factors, exclusive of evidentiary weight. Aggravated assaults and robberies present fewer challenges in terms of better matching crime severity than do homicides and rapes. Further, the nature of homicides and rapes may elicit

Exclusions

Certain defendant demographic data is excluded from the vignettes. Specifically, race and ethnicity are not included. Since prior research has demonstrated that race and ethnicity are extralegal sentencing factors that may impact judicial sentencing decisions (Albonetti, 1997; Everett & Wojewicz, 2002; Steffensmeir & Demuth, 2000), excluding them from the vignette controls for that influence. Further, judges participating in the study – possessing a general understanding that this is a qualitative study of factors that may influence judicial sentencing decisions – may be guarded with their analysis if a vignette contains information on race and ethnicity. Indeed, there is considerable risk that they may feel compelled to provide socially acceptable answers. Additionally, the defendant's age is excluded. Depending on context, age may be legally relevant at sentencing or it may be extralegal.

Finally, judges were informed that all four defendants are male. Since some studies have found that gender plays an extralegal role in sentencing decisions (Mustard, 2001; Rodriguez, Curry & Lee, 2006), controlling for this potential influence can strengthen the effectiveness of the vignettes. As opposed to race, ethnicity, and age, excluding gender from the vignettes is practically challenging. In conversations following the imposition of sentence by the judges, avoiding all pronouns when referring to the defendant is not feasible. Male was selected as the gender of the defendant because the vast majority of defendants appearing before criminal court judges are male.

greater emotional responses that influence sentencing determinations, thereby interfering with the ability of the vignettes to isolate the influence of evidentiary weight.

Benefits and Limitations

The use of vignettes can facilitate the interviewing process in several ways. First, since vignettes are stories that do not reference actual events, they can provide an opportunity for interviewees to discuss related issues in a non-threatening way (Hughes, 1998). In fact, as opposed to asking judges about actual cases, the presentation of fictitious hypothetical cases is a less threatening and more relaxed way of exploring the topic and may yield more complete and unfiltered responses.

Further, vignettes may serve as ice breakers by prompting judges to recall actual cases when they compare their personal experiences to those presented in the vignette (Hughes, 1998). Since N.Y.S. Supreme Court judges deal with a high volume of criminal cases – some sharing similarities with the vignettes presented – the interview may naturally progress from judges comparing their analysis and decision making process in theory to discussing actual cases they adjudicated in practice.

The ability to selectively choose the contents of the vignettes is extremely beneficial. By creating examples that vary predominantly in evidentiary type and strength, the influence of evidentiary weight on sentencing can be isolated. Moreover, the interview process allowed judges to explain their rationale immediately after rendering a hypothetical sentence, thus capturing more accurate descriptions since their reasoning is freshest at that time.

The use of case scenarios can be helpful for other reasons as well. Some judges may be unaware that evidentiary factors are influencing their decisions. Their responses to hypothetical case scenarios may be less guarded than responses to actual cases and, therefore, provide insight that might otherwise be missed.

There are, nevertheless, several limitations with using hypothetical case scenarios. A common concern when employing vignettes as a research tool “surrounds the distance between the vignette and social reality, what people believe they would do in a given situation is not necessarily how they would behave in actuality” (Barter & Renold, 2000, p. 311; see, also Hughes, 1998). While some prior research has found that vignettes “mirror how individuals act in reality” others have argued that there is insufficient information to draw parallels between the vignettes and reality (Barter & Renold, 2000, p. 311; Hughes, 1998). For purposes of this study, hypotheticals can be instrumental not only in studying how judges would act in a particular situation, but also for understanding the ways in which judges think about and perceive different evidentiary forms. Barter and Renold (2000, p. 312) argue that “it is not the outcome (or action) that is of research interest, for this will always be situationally specific, but the process of meanings and interpretations used in reaching the outcome that is of central concern to social scientists.” Additionally, for purposes of this study, the combined use of mixed methods and a multi-method qualitative approach helps to capture the social reality by corroborating or contradicting findings unearthed during the vignette section of the qualitative interviews.

A second concern is that the judges may try to give what they believe to be socially desirable responses (Hughes, 1998; Barter & Renold, 2000). Here, this concern is minimized. First, given the huge range of potential sentences a judge may legally impose for robbery and assault in the first degree under N.Y. law, judges were instructed to impose a sentence within the designated range established by statute. By directing judges

to legally sanctioned sentences, judges should be less concerned about socially desirable responses.

Another limitation is that even the most detailed and descriptive presentation of a case will, naturally, miss nuances and subtleties that appear in an actual trial (where judges see the witnesses and are able to evaluate their credibility, as well as examine the quality of the forensic evidence). This limitation notwithstanding, the presentation of hypothetical case scenarios can provide key data instrumental in studying judicial thought processes involved in sentencing determinations.

Finally, the participation rate for the vignette portion of the interview was approximately 54% (n=21) (see footnote 31). While full judicial participation would have been ideal, the data collected from the 21 participants provides rich information; not only in the actual sentence lengths imposed but also in the accompanying explanations and insights.

Data Analysis

All audiotaped interviews (n=30) were transcribed verbatim. Handwritten notes taken during the remaining interviews (n=11) contain direct quotes and were typed verbatim. Following transcription, data collected during the qualitative interviews were coded. Coding was performed in order to “fracture the data and rearrange them into categories that facilitate comparison between things in the same category and that aid in the development of theoretical concepts” (Maxwell, 2005, p. 96).

Initially, data was sorted into “organizational categories” (Maxwell, 2005, p. 97) that “function[ed] primarily as bins for sorting data for future analysis” (Maxwell, 2005: 97). Given the large volume of transcribed pages (approximately 1200 pages), data was

initially divided according to broad subject categories (e.g. perceptions of evidentiary forms, views on juries, judicial discretion, sentencing factors). Each of these broad categories was then subdivided into smaller categories. For example, data collected regarding judicial perceptions of different evidentiary forms were initially sorted into categories based on the type of evidence (i.e. forensic or witness based). Each of these categories was then fractured into subcategories. For example, forensic evidence was divided into categories such as DNA, fingerprints, ballistics and chemists while witness evidence was divided into the categories of eyewitnesses, character witnesses, alibi witnesses, witness credibility and witness reliability. Data was then further fractured based on the judge's response for ease in future analysis. For example, in the category of alibi witnesses, data was coded into the following categories based on the judges' responses: Alibi testimony is powerful, alibi testimony usually hurts the defense, strength of alibi testimony is dependent on the relationship, and alibi witnesses are irrelevant.

An inductive, grounded theory approach was used to develop theoretical categories that "place the data into a more general or abstract framework" (Maxwell, 2005, p. 97). Themes emerged in comparing the different categories of evidence (e.g. judges are more confident in guilt if the case contains forensic evidence; single witness identifications raise reasonable doubt). Several emerging themes were unexpected (e.g. judges expressed confidence in the scientific processes involved with DNA evidence, yet many of them do not understand the processes involved; the "showmanship" -- rather than the expertise and demonstrated skill -- of forensic expert witnesses is a primary judicial focus in evaluating the quality of the evidence). Data collected from each new

case was compared with data from prior cases and emerging themes and hypotheses were consistently refined to accurately reflect the complete body of data accumulated.

CHAPTER 4. QUANTITATIVE ANALYSES

This chapter describes the results and implications of quantitative analyses performed to examine whether and to what degree evidentiary quantity and type influence judicial sentencing decisions. Studying whether this influence exists, and to what extent, is a critical first step toward understanding how evidentiary weight might act as an extralegal sentencing factor for judges. I report the results of a series of regression models. After exploring the influence of legally relevant sentencing determinants on violent felony (i.e. homicide, aggravated assault, robbery and rape) trial and plea cases (n=483) in Model 1, the second regression model examines the influence of evidentiary quantity and type on these cases. The results demonstrate that both quantity of physical evidence and evidentiary type influence sentence length. These findings are described below.

While informative, the results observed may be capturing factors that are not wholly indicative of judicial sentencing discretion. This is because decisions to plead guilty and sentencing decisions in plea cases are determined concurrently, consider the probability of conviction (Bushway, Redlich & Norris, 2014), and are largely subject to prosecutorial discretion. In contrast, sentencing post-conviction is a distinct phase in trial cases, occurring after all issues of guilt have been adjudicated. It is here that judges have the central role. Consequently, sentencing decisions in trial cases are a purer reflection of judicial discretion (Johnson, 2003). Further, trial cases provide judges with greater exposure to the guilt-phase evidence than do plea cases. Therefore, the results of a third model, utilized to explore the role of evidentiary weight at sentencing exclusively in trial cases, are reported and discussed (n=157) below.³⁸

³⁸ A plea only model is provided in Table 4.5. Since sentencing post-plea is a greater reflection of prosecutorial discretion and is often determined during plea negotiations, excluding these cases permits a

As indicated in Chapter 3, the following hypotheses are examined with data from a public-use dataset entitled *Impact of Forensic Evidence on the Criminal Justice Process in Five Sites in the United States, 2003-2006* (Peterson & Sommers, 2010).³⁹

H1: Violent felony cases that contain more pieces of physical evidence will result in longer prison sentences.

H2: When the guilt-phase evidence includes a forensic laboratory report, the defendant convicted at trial will be sentenced to a longer period of incarceration.

H3: When the guilt-phase evidence includes one or more eyewitnesses, the defendant convicted at trial will be sentenced to a longer period of incarceration.

H4: When the guilt-phase evidence contains more pieces of physical evidence, the defendant convicted at trial will be sentenced to a longer period of incarceration.

Ordinary Least Squares regression models are employed to predict sentence length. In Model 2, I utilize the full sub-sample of criminal cases to test the influence of the quantity of physical evidence as well as the relative influences of eyewitness evidence and forensic evidence on judicial determinations of sentence length. I then conduct a set of analyses where I distinguish the predictors of sentence length exclusively in trial cases. Diagnostics reveal that no significant collinearity problems exist in the analyses reported below. Specifically, the highest VIF in Model 1 is 2.917 (tolerance = .343) for assault with injury, and the highest VIF in Model's 2 and 3 is 3.904 (tolerance = .256) for robbery without injury.⁴⁰

more accurate analysis of the role of evidentiary weight on **judicial** sentencing discretion. As Table 4.5 shows, the only significant evidentiary variable in plea cases was the presence of a lab report, which resulted in a 53.57% increase in sentence length.

³⁹ These data are available at <http://doi.org/10.3886/ICPSR29203.v1>. For a description of the data and the sub-sample of data utilized for this study, see Chapter 3.

⁴⁰ IBM SPSS v. 23 multiple imputation procedures were utilized to impute missing data on the number of prior convictions variable. Estimates from the pooled imputed data are used in the models. Since SPSS does not report pooled data for VIF, the numbers reported are based on the imputed model with the highest VIF.

A. Description of Variables

In Chapter 3, a brief description of the variables utilized in the regression models is provided. In the section below, a more detailed description of the operationalization of all variables is discussed.

i. Dependent Variable

The dependent variable is an overdispersed count variable measuring the number of months to which the defendant was sentenced. As shown in Table 4.1, defendants were sentenced to an average of 166.57 months of incarceration (or approximately 14 years), although sentences ranged dramatically from 1 to 2,760 months (s.d. = 300.28). After taking the natural log of sentence length, its distribution approximates a normal curve.

[TABLE 4.1 AND FIGURES 4.1 AND 4.2 ABOUT HERE]

ii. Independent variables

Three categories of independent variables are included in these models: legally relevant sentencing criteria, evidentiary variables, and control variables. Each of these is discussed, in turn, below.

Legally Relevant Sentencing Criteria

Three variables reflect legally relevant sentencing characteristics; these include offense seriousness, the defendants' criminal history, and mode of conviction. Crime type, one measure of offense seriousness, is legally relevant for sentencing purposes and has always been used to delineate appropriate sentencing ranges under both discretionary

and mandatory sentencing regimes, and victim injury or harm caused by the offense is another indicator of offense severity (Robinson, 2008). Together, they provide a strong indicator of offense seriousness. Therefore, crime type is combined with victim injury (measured by whether or not the victim received medical treatment) to operationalize *offense seriousness*. More than one quarter of the convictions are homicide cases in these data (26.0%; n=126, see Table 4.1). The remaining cases are distinguished as follows: 7.4 percent are aggravated assaults in which the victim received medical treatment (n=36); 27.3 percent are aggravated assaults in which the victim did not receive medical treatment (n=132); 11.8 percent are rapes in which the victim received medical treatment (n=57); 1.4 percent are rapes in which the victim did not receive medical treatment (n=7); 2.9 percent are robberies in which the victim received medical treatment (n=14); and 23.1 percent are robberies in which the victim did not receive medical treatment (n=112).

The *number of prior convictions* in a defendant's criminal history is included in the models. As shown in Table 4.1, defendants had an average of 5.2 prior convictions, although criminal conviction histories ranged from 0 to 78 (s.d.=6.61).⁴¹ Mode of conviction distinguishes cases disposed via *plea* agreement from convictions at trial. In the interests of efficiency, plea bargains are often associated with shorter sentences because consideration at sentencing provides an incentive to plea (Feeley, 1979; but see

⁴¹ Unfortunately, the data do not distinguish criminal history by the type, level of seriousness, or punishments associated with previous offenses. Nevertheless, this basic control for criminal history makes it possible to partial out at least some portion of the sentencing decision that is related to the risk posed by the defendant. In total, there are 95 missing values on prior convictions; as indicated in footnote 40, IBM SPSS v. 23 multiple imputation procedures were utilized to impute missing data.

Bibas, 2004). Approximately two-thirds of these convictions are resolved through a plea agreement (67.4%; n=326) and the remaining third are decided at trial.⁴²

Evidentiary Weight

Three evidentiary variables are included in the model: (1) *presence of a forensic lab report*, (2) *quantity of physical evidence*, and (3) *presence of eyewitnesses*. The measure of lab-examined forensic evidence captures the existence of various types of crime scene evidence examined in the case, such as firearms, latent prints, drugs, biological evidence (primarily blood), pattern evidence, generic objects, and natural and synthetic materials (Peterson et al., 2010); 41.9 percent of the cases (n=203) involved the examination of physical evidence in a crime lab.

The quantity of evidence is a sum of the amount of physical evidence collected in a case and submitted to the crime laboratory for analysis (including biological evidence, ballistic evidence, tangible evidence, drugs, blood, DNA, fingerprint, clothing, hair, tissue, fibers, gun, trace, cartridges and other unspecified physical evidence). On average, 1.19 pieces of physical evidence were collected and submitted for analysis in these cases (s.d.=1.75; range=0-8).

Finally, the presentation of eyewitness testimony may serve as an influential, corroborating source of information that can strengthen the prosecution's case. Cases with no eyewitnesses are distinguished from cases with any eyewitnesses; in total,

⁴² While the percentage of trial cases in the sub-sample is a little less than a third, the cases in the original data set, including cases disposed of at earlier stages in the process and burglary cases (n=4205), contains 379 trial cases (9%). The large percentage of trial cases in the sub-sample may be due to the serious and violent nature of these cases and the failure of the cases to resolve at earlier stages of the process, when plea offers are often more favorable.

slightly more than half of the cases (57.4%; n=278) include eyewitnesses as part of the evidentiary package.

Control variables

Much research has found sentencing disparities by race that penalize African American and Latino defendants relative to white defendants (Feldmeyer & Ulmer, 2011; Johnson et al., 2011), and male defendants relative to female defendants (Mustard, 2001). Consequently, defendant *race* and *sex* are controlled in a series of dummy variables: Suspect Black, Suspect Latino (Hispanic is treated as a mutually exclusive racial category in these data), Suspect White, and Suspect Asian. Any defendant not coded as 1 in one of these four groups must be in a different race/ethnicity category or unknown. In these analyses, I include dummy variables for African American (48.1%; n=233) and Latino defendants (12.6%; n=61), respectively, because these two groups have been historically subject to more punitive sentences, when legal characteristics of cases are held constant (Albonetti, 1997; Everett & Wojtkiewicz, 2002). Consequently, the reference category includes Whites (16.3%; n=79), Asians (.4%; n=2), defendants of other races/ethnicities, and those whose race/ethnicity is unknown (22.5%; n=109). In addition, the vast majority of defendants are male (92.6%; n=448). Victim race and gender are also controlled (Baldus, Pulaski, & Woodworth, 1983; Baumer, Messner & Felson, 2000). In total, one third of the victims are African American (33.3%; n=161) and 13.4 percent are Latino (n=65). The reference category includes White victims (26.9%; n=130), Asian victims (1.9%; n=9), victims who may be of another race/ethnicity, and those whose race is unknown (24.6%; n=119). Nearly two-thirds of the cases involve male victims (61.2%; n=296). Finally, a variable distinguishing the *site* in which the cases are processed to

control for systematic differences in the legal or procedural treatment of cases across state jurisdictions is included in the models. Specifically, Los Angeles (38.2%; n=185) is distinguished from all Indiana locations (61.8%; n=299).⁴³

B. Results

Before examining the multivariate models regressing custodial sentence length on the various legal, evidentiary, and control variables, the extent to which the type and quantity of evidence varies across the four violent crime categories is first explored. Table 4.2 shows that, on average, cases in which defendants are convicted of homicide have a larger and more variable quantity of physical evidence ($\bar{x} = 3.14$) compared to rape ($\bar{x} = 1.25$), robbery ($\bar{x} = 0.32$), and aggravated assault ($\bar{x} = 0.36$) cases. As might be anticipated, a significantly larger proportion of homicide convictions (86.5%) include lab-examined forensic evidence at the guilt phase, followed by rape (48.4%), robbery (25.4%), and aggravated assault (18.5%). Finally, a significantly larger proportion of both aggravated assault (66.7%) and homicide (72.2%) convictions are cases that involve one or more eyewitnesses compared to robbery (47.6%) and rape (23.4%) convictions.⁴⁴

[TABLE 4.2 ABOUT HERE]

⁴³ All five of the jurisdictions in these data apply either the California (Los Angeles site) or Indiana Penal Codes (Indianapolis, South Bend, Evansville and Fort Wayne sites). While a comparison of these penal codes shows minor distinctions, the general substance of the crimes included in this study are substantively analogous. For example, both states contain virtually identical elements for the crime of robbery (i.e. requiring that the defendant take property from another by force or fear). Likewise, homicides in both states require the killing of another accompanied by an applicable mens rea (knowing, intentional or willful). The offense categories are thus reasonably consistent across jurisdictions. Minor differences in sentence ranges exist between these jurisdictions. For example, Indiana's sentencing range for a Class 5 Robbery is between 1 and 6 years (as of 2014; prior to that time, the sentence range was 2 to 8 years), whereas the possible sentence length in California for Second Degree Robbery (with similar elements) ranges between 2 and 5 years. To control for these variations, Los Angeles and Indiana jurisdictions are distinguished in the model.

⁴⁴ Recall that the eyewitness variable includes all eyewitnesses (including victims), which helps to explain the fact that nearly one quarter of rape cases involve eyewitnesses.

The OLS regression models are based on all homicide, rape, aggravated assault, and robbery convictions resulting in custodial sentences in the sample. For ease of interpretation, the estimated coefficients are exponentiated to calculate the percentage change in number of months sentenced for each independent variable. Since the natural log of sentence length is used in this analysis, percent change in number of months sentenced is calculated according to the formula:

$$(e^b - 1) * 100 \quad [\text{Eq. 1}]$$

Model 1 in Table 4.3 illustrates the effects of legally relevant sentencing criteria on sentence length. The results largely confirm expectations. Specifically, compared to homicide cases, defendants convicted of each of the other types of crime (i.e. aggravated assault with or without medical treatment, robbery with or without medical treatment, and rape with or without medical treatment) are, on average, sentenced to significantly shorter periods of incarceration. Further, those cases resolved via plea involved, in general, 43.5 percent shorter custodial sentences compared to trial convictions. Finally, while the coefficient for number of prior convictions is in the expected direction, it is statistically insignificant.

[TABLE 4.3 ABOUT HERE]

In Model 2, evidentiary weight is incorporated, including the presence of forensic evidence, the presence of eyewitnesses, and the quantity of physical evidence in the case, as well as all of the control variables. In terms of evidence type, cases containing a forensic lab report, on average, result in a 51.4 percent increase in number of months sentenced compared to cases without such evidence ($p < .01$), whereas the presence of eyewitness evidence is not a significant predictor of sentence length. The quantity of

physical evidence in the case positively and significantly influences the length of the custodial sentence, such that each additional piece of physical evidence results in a 12.5 percent increase in number of months sentenced ($p < .05$), consistent with Hypothesis 1. These results suggest that sentencing decisions are influenced not only by the type of evidence available in a case (i.e. forensic evidence vs. witness-based evidence), but also by the total amount of physical evidence marshaled against the defendant, even controlling for offense severity and other relevant factors at sentencing.

The remainder of the findings in Model 2 are largely in the expected direction with the exception of victim and defendant sociodemographic characteristics, which are insignificant. Only mode of conviction (plea), offense seriousness, and site are found to significantly influence sentence length beyond evidentiary weight. Defendants who plead guilty in violent felony cases are sentenced, on average, to 45.4 percent shorter periods of incarceration compared to those who are found guilty at trial, all else equal ($p < .001$). In general, compared to those convicted of homicide, the custodial sentence imposed for defendants in other violent felonies is between 51.7 (rape with injury) and 91.2 percent (aggravated assault without injury) shorter. Finally, cases in Los Angeles courts are associated with shorter sentences than those in the Indiana jurisdictions, controlling for known legal and extralegal sentencing factors ($p < .01$).

Turning to post-conviction sentencing in trial cases, the model estimated in Table 4.4 predicts sentence length exclusively for defendants convicted at trial ($n=157$).⁴⁵ The results can therefore be interpreted as a purer reflection of judicial sentencing decisions post-conviction, as plea negotiations in the shadow of the trial are removed. Trial

⁴⁵ A post hoc power analysis reveals that the trial model ($n=157$) achieved power of 1.000 with an alpha level of .01.

convictions that involve a forensic lab report during the guilt phase result in an 82.2 percent increase in number of months sentenced compared to cases without lab reports ($p < .05$). Further, each additional piece of physical evidence results in a 16.2 percent increase in the number of months sentenced ($p < .05$). These results provide strong support for Hypotheses 2 and 4. Contrary to Hypothesis 3, the effect of eyewitnesses on sentence length in trial cases is insignificant, just as it is in the full model (see Table 4.3).

[TABLE 4.4 ABOUT HERE]

Victim and defendant sex and race continue to show null effects and cases in Los Angeles courts are associated with shorter sentences than cases in Indiana jurisdictions. Further, all offense seriousness variables indicate a statistically significant decrease in number of months sentenced compared to homicide cases, with the exception of rape cases with medical injuries. Notably (and in contrast to prior models), the defendant's criminal history has a small but statistically significant effect in convictions at trial, in that each additional prior conviction results in a 3.36 percent increase in the number of months sentenced ($p < .05$).

C. Discussion

The extralegal influences of race, gender, and other ascribed characteristics on the sentences imposed by judges are well established, and a majority of scholarship in this area finds serious and troubling disparities in the punishments given to different types of offenders in similar kinds of cases. These quantitative analyses contribute to that

literature by investigating additional sources of extralegal sentencing disparities, in the form of the quantity and type of evidence presented during the guilt phase of violent felony trials. In the end, both the presence of forensic evidence and the quantity of physical evidence influence sentence length for cases disposed via trial. These findings provide cause for concern because the type and quantity of evidence used to establish the elements of the crime and obtain the conviction should have no bearing on the severity of sentences imposed by judges once offense severity is considered. In other words, while the nature and scope of evidence amassed in criminal cases is imperative for bringing offenders to justice, the type and quantity of guilt phase evidence loses legal relevance at the point of conviction. This is because sentencing is not a time to re-evaluate whether the defendant should have been convicted during the guilt phase but, instead, a time to determine the appropriate sentence to impose on a defendant already adjudicated guilty.

When judges use evidentiary strength (i.e. weight of the evidence relevant to establishing the elements of the crime) as a basis for their decisions in the sentencing phase, they are in effect revisiting the guilt-phase decisions of the trier of fact and are thus overstepping to evaluate the merit of the case *after* it has already been decided. This undermines the “finality” of jury decisions, a sacred legal doctrine. Yet trial judges are shown to be more punitive with defendants when the evidentiary package includes more pieces of physical evidence and when at least some of that evidence is subject to forensic examination in crime labs. Alternatively, cases without forensic evidence and cases containing little or no physical evidence are associated with shorter sentences, perhaps as a means of partially mitigating perceived guilt-phase injustices. Indeed, a possible explanation for the observed differences in the influence of type and quantity of guilt-

phase evidence at sentencing is that judicial sentencing decisions may be motivated at least in part by a judge's confidence in the accuracy of the verdict.

The findings regarding forensic evidence and witness-based evidence, respectively, help to inform the manner in which evidentiary weight acts on discretionary judicial decisions. Lab-examined forensic evidence refers to a wide array of physical evidence including firearms, drugs, latent prints, generic objects, biological evidence (primarily blood), natural and synthetic materials, and pattern evidence (Peterson et al., 2010). Forensic evidence is recovered by the police and brought to the crime lab for testing, and care is taken to preserve the chain of custody and ensure that the item being tested has not been contaminated (Peterson et al., 2010). Once in the lab, trained personnel, experienced in forensic testing, use standardized procedures to conduct their analyses (even if some of these procedures rest on shaky scientific foundations) (Saks & Koehler, 2005). In the vast majority of cases, police officers, detectives and lab personnel who are involved in recovering, transporting and testing physical evidence have no personal involvement in the case. The results of forensic analyses are evaluated using ostensibly objective criteria.

To the extent that lab reports represent objective evidence, forensics may bolster judicial confidence in a defendant's guilt. This, in turn, serves as an influential factor to judges vested with wide discretionary powers in determining appropriate sentences for convicted defendants. For example, a defendant may be determined to have been present at the crime scene by an eyewitness or by the recovery of his fingerprints from the location. A judge may feel more confident in guilt when presented with fingerprint evidence, in contrast to the subjectivity inherent in witness testimony. While the

defendant's culpability, as a matter of law, is the same regardless of the evidentiary form that the prosecutor presented in the case, increased confidence in guilt may facilitate the imposition of harsher sentences in violent offenses. Or, inversely, seemingly less reliable subjective evidence at trial may undermine the judge's confidence in the finding of guilt and encourage him or her to soften the punishment for cases in which he or she is more uncertain.

The existence of a null association between witness-based evidence and sentence length suggests that this evidentiary form has no discernible effect on the post-conviction confidence of judges regarding the guilt of defendants. Compared to forensic evidence, witness-based evidence is less objective and subject to greater credibility and reliability concerns. Determining credibility is a complex process that involves considerations about whether the witness has a stake in the outcome of a case, his or her trustworthiness, and the relationships between the witness, the defendant, and the victim (Loftus, 1974; Shermer, Rose & Hoffman, 2011). Likewise, the background (e.g., criminal history) and perceived competence a witness shows on the stand can influence whether the jury or judge believes the testimony (Shermer, Rose & Hoffman, 2011).

In addition to credibility, the reliability of a witness' testimony may be questioned. Specifically, factors external to the truthful intent of the witness may cast doubt on the accuracy of the testimony; these include concerns about the witness' memory, lighting conditions, the distance from which the incident was observed and the victim's eyesight (Sanders, 1984; O'Neill Shermer, Rose & Hoffman, 2011). In these situations, even witnesses who are deemed truthful may be disbelieved. For these reasons, much eyewitness identification testimony has been found to be inherently

unreliable (O'Neill Shermer, Rose & Hoffman, 2011; Sharps, Hess, Casner, Ranes & Jones, 2007; Sanders, 1984)).

While eyewitnesses to a crime can provide strong, corroborating support in cases where the trier of fact determines that the witness is credible and reliable, the nature of witness testimony necessitates a subjective assessment of both credibility and reliability that requires individualized judgment calls on matters that are often complex and ambiguous. This inherent uncertainty may serve to weaken the strength of the evidentiary package and reduce the confidence level the sentencing judge possesses regarding the defendant's guilt. Despite the fact that sentencing judges do not have a fact-finding role regarding guilt at sentencing, judges (consciously or subconsciously) may have a difficult time separating their feelings regarding the defendant's guilt from their sentencing roles in the case.

The results of the trial only model are important for several reasons. In the post-trial arena, sentencing decisions are not negotiated or predetermined (except within broad sentencing guideline ranges) but, instead, are the primary domain of the trial judge and therefore reflect genuine processes of judicial discretion. As opposed to pleas, trials provide judges with the greatest opportunity to observe and evaluate all of the evidence presented in a case. A presiding trial judge is in the position to hear the words and observe the demeanor of witnesses' first-hand, to view and examine physical evidence presented at trial, and to evaluate the testimony of forensic experts. Moreover, the trial judge may observe evidence which he or she ultimately deems to be inadmissible at a pre-trial hearing – evidence that even juries do not have the opportunity to observe. This higher degree of exposure to the evidence enables judges to form opinions about the

strengths and weaknesses of different types of evidence and to evaluate the general quality of evidence presented in specific cases. Arguably, the penalty in sentence length associated with evidentiary strength for defendants convicted at trial is reflective of increased judicial confidence in guilt for two reasons: first, exposure to multiple pieces of physical evidence provides mounting corroboration of guilt and, second, forensic evidence is perceived to be more objective and less impeachable than are other kinds of evidence.

In the end, the results of the quantitative analyses indicate that the weight of the evidence in a case has a strong influence on sentence severity. While these results establish that there is an association between evidence type and quantity and sentence length and that this association is significant, the reasons for this association cannot be definitively answered from these quantitative analyses alone. Said differently, the only way to understand why judges sentence defendants based on evidentiary weight is to speak directly with sentencing judges. Understanding judges' perceptions of different evidentiary forms, their views regarding jury verdicts and their finality, as well as their perceptions of their own roles in sentencing in light of these issues is a critical second step that must be completed in order to fully explore this issue. Only through direct dialogue with sentencing judges can an accurate assessment be made of the role of evidentiary strength on judicial sentencing decisions.

CHAPTER 5: JUDICIAL PERCEPTIONS OF EVIDENTIARY WEIGHT

The weight of the evidence present in a criminal case is a critical factor in determining the guilt or innocence of the accused defendant. Evidentiary packages may consist of various forms and quantities of forensic and other non-scientific witness-based evidence. As recent studies demonstrate (Peterson et al., 1987, 2010, 2013) and Chapter 4 of this dissertation confirms, evidentiary type continues to play a role during the sentencing phase of a criminal prosecution. Specifically, all else equal, cases with forensic evidence result in longer sentences than cases without forensic evidence. The findings presented in Chapter 4 further reveal that the quantity of evidence in a case influences sentence length, in that additional pieces of physical evidence result in longer sentences for convicted defendants, and that type and quantity of evidence influence judicial sentencing decisions post-verdict in trial cases. In these data, eyewitness evidence is found not to be influential at sentencing.

While the existence and prevalence of this influence was explored using quantitative analyses in Chapter 4, the question of why these influences exist cannot be answered using quantitative methods alone. Therefore, this dissertation further examines these questions through in-depth interviews with sentencing state court judges in New York. In order to better understand why these evidentiary influences exist, it is essential to first study how judges perceive different evidentiary forms and packages. This preliminary analysis is a critical first step to understanding why and how the weight of evidence affects judicial sentencing determinations post-verdict.

In order to explore these issues, this chapter will report findings regarding judicial perceptions of various forensic and witness-based evidentiary forms. Within each broad

category (i.e. forensic and witness-based), different types of evidence exist. For example, among other evidence, forms of forensic evidence include DNA, fingerprint, ballistic, and chemist evidence. In the arena of non-scientific witness testimony, different types of witness-based evidence may be presented. These may include eyewitnesses, character witnesses, and alibi witness, among others. Further, the testimony of expert psychologists and psychiatrists, which has been characterized by some as a “soft” form of scientific evidence, often plays a key role in criminal prosecutions as well.⁴⁶

In order to unpack distinctions in judicial perceptions of various evidentiary forms and the evidentiary weights or values attributed to them, different types of evidence, within each category, will be explored individually. By fragmenting the analyses in this way, distinctions between various pieces of evidence and their unique characteristics can be more accurately observed and judicial perceptions of evidentiary weight more thoroughly understood.

To this end, this chapter is organized as follows. First, judicial perceptions of DNA, fingerprint, ballistics and chemist evidence will be addressed individually, in turn. At times, comparisons will be made between different forensic evidentiary forms including perceptions as to the weight of distinct types of evidence. I will then address judicial perceptions regarding the absence of forensic evidence in a case. Following the analyses regarding forensic evidence, I will explore judicial perceptions regarding non-scientific witness-based evidence as follows. First, I will individually explore judicial

⁴⁶ While a psychologist may be classified as an expert witness, most psychological testimony is fact-based in that the testimony often regards “matters that [the psychologist] has perceived through the senses” (Gutheil, 2003, p. 385). Additionally, psychologists may also testify to diagnoses, consequences or treatment plans (Gutheil, 2003). Given the fact based nature of this testimony, psychologists are discussed in the witness-based evidence section of this paper.

perceptions of four types of witness evidence: eyewitnesses, character witnesses, alibi witnesses, and psychologists. I will then discuss the various factors judges consider when determining the credibility and reliability of witnesses. These factors are important because they provide insight into the ways that judges evaluate the strengths and weaknesses of witness-based evidence. Finally, I will explore judicial comparisons of the value of forensic evidence, viewed as a whole, compared to witness-based evidence.

A. Judicial Perceptions of Forensic Evidence

As I will argue below, while judges perceive forensic evidence to be a stronger and more objective form of evidence than non-scientific witness-based evidence, judicial perceptions regarding the probative value and strength of forensic evidence vary based on the type of forensic evidence presented and case characteristics, among other factors. First, I will discuss the perceptions of judges regarding DNA, fingerprint, ballistics, and chemist evidence individually. I will then turn to judicial perceptions about case strength in the absence of forensic evidence.

Judicial Perceptions of DNA Evidence

On the whole, the vast majority of judges view DNA evidence as the strongest and most objective form of evidence available. While more than half of the judges discussed the CSI effect and the overreliance of juries on DNA evidence (for a discussion of the CSI effect, see Cole & Dioso-Villa, 2011; Hayes-Smith & Levett, 2011; Hughes & Magers, 2007; Kim, Barak & Shelton, 2009; Shelton, Barak & Kim, 2011), most expressed their own desires to have DNA evidence included as a part of the evidentiary

package: “I always prefer to have certainty and you have that with DNA evidence. For truth-seeking courts, it is always best to be sure.”

While approximately a quarter of the judges discussed the importance of determining the chain of custody of the sample (e.g. blood) tested and ensuring that contamination did not occur, consistent with scholarly literature, the science of DNA evidence and the procedures utilized to test the evidence were largely accepted without question (Saks & Faigman, 2008). Instead, the focus in discussing DNA evidence was on the ability of the DNA expert witness to testify clearly, generate interest, and the overall showmanship and likability of the witness. Almost a quarter of the judges conceded that they lack an understanding of “how DNA evidence works” as well as the procedures utilized by testifying experts (for similar conclusions see McQuiston-Surrett & Saks, 2009; Saks & Faigman, 2008; Wojcikiewicz, 2013). Moreover, consistent with prior findings, more than half noted that many attorneys do not have the ability to effectively question DNA experts because they do not possess an understanding of the scientific principles and mechanics involved (see Edmond & Roque, 2012; Saks and Faigman, 2008). Finally, while judges expressed their confidence in DNA evidence, some emphasized that it is merely one piece of the evidentiary package and rarely establishes all of the required elements of the crime. In this section, I will discuss each of these themes.

First, judges expressed overwhelming confidence in DNA evidence. One stated that “DNA is the golden child . . . it is close to infallible.” Another explained: “I will always trust DNA because I trust scientists.” Yet a third expressed his perception that DNA evidence is ‘bulletproof’: “DNA is 100% reliable . . . It comes back so many billion to

one so it's the most reliable form of evidence. There is no problem with perjury, false confessions, misidentifications – it eliminates all of that type of thing.” In a similar expression of confidence in the certainty of DNA evidence, another judge stated: “When you get an expert up there who says that ‘with 99.7% certainty they can rule out A, B and C,’ you have to give it weight.” One summed up his perception of the strength of DNA evidence in one sentence: “If the issue is identification, it is a rock crusher.”

While the vast majority view DNA evidence as an extremely valuable and probative form of evidence, a majority of judges are critical of the jury's demand for and blind acceptance of forensic evidence in general and DNA evidence in particular:

I think, unfortunately, people expect [to see] DNA and other forensic evidence because they watch too much television. I think that DNA analysis is really boring. I think that jurors totally turn off to it. All they need to hear is that it is DNA and that it is the defendant's DNA and the case is decided.

Another stated, “You have to almost dissuade them and say, “Okay, that's TV and this is real life. Try to keep those two separate.” A third expressed the challenge in managing the CSI Effect: “The system is now inundated with gratuitous requests for DNA simply because there is no choice but to order it so as not to face the question of ‘did you test for DNA?’ even though it is not necessary if the jury used its common sense and evaluated credibility of witnesses independent of forensic evidence.”

According to one judge,

DNA evidence is overwhelmingly reliable and it is actually a terrible problem. Juries are not always people who really grasp many things. One of the great drawbacks where I sit is that DNA has been almost automatic and mandatory.

Some expressed concerns that defense attorneys often do not possess the skills and knowledge to effectively cross-examine expert DNA witnesses. One judge stated: “I

haven't seen anybody lay much of a hand on a DNA witness. I don't think that I have even heard of it." A second stated: "DNA people almost always get a free pass because lawyers are intimidated by DNA and do not know how to attack it." Finally, a handful expressed frustration at this lack of skill: "The adversarial system fails its purpose when the attorney is not skilled enough to challenge the expert" (for similar conclusions see Edmond & Roque, 2012; Saks & Faigman, 2008).

Judges stated different reasons for this lack of skill. Some point to the fact that most attorneys are not trained to understand DNA evidence. "There have been very few cases where anybody has been able to show that there was an error committed in the DNA procedure. I think it is because it's a complicated area that unless they get specific training on it, they don't know the answers or the questions to ask to get the answers that they are looking for." Another noted that training is available but that attorneys do not always take advantage of the training:

In a case with DNA, both sides know way in advance that there is going to be DNA evidence. They have an opportunity to, if they are not really up on it, to take the time to learn what they need to ask. There are tools out there I think for both sides, to go to seminars or to read materials on how to effectively question or cross-examine a DNA expert. Most don't bother.

Others argued that even in situations where the defense attorney is knowledgeable enough to effectively cross-examine a DNA expert, jurors usually "tune out" because DNA testimony is "highly technical" and "boring."

Several judges distinguished by "attorney type" in evaluating the competency of defense counsel to cross examine a prosecution DNA witness: "In high profile cases, where an individual has the ability to hire their own counsel, it could be good. Really a pleasure to watch. If you are dealing with a public defender who has been doing this

perhaps too long, the cross examination is useless –it’s actually really sad to watch.” In the end, almost a quarter of those interviewed conceded that even they did not understand the science behind DNA evidence or the specific mechanics involved (see McQuiston-Surrett & Saks, 2009; Saks & Faigman, 2008; Wojcikiewicz, 2013).

The perception that DNA evidence is simply too “highly regarded” to attack, regardless of attorney skill, was articulated by a handful of judges:

I think that the science is pretty well established and has been through the fire test. There are some new DNA technologies that they are now reviewing in the state. I find that it’s not a lack of skills or knowledge on the part of the defense attorney’s cross-examination. It’s more of a lack of ability to attack the science of it because it is so well accepted in our society.

Another echoed a similar sentiment: “The protocols are simply too strong to attack.”

A number of judges concluded that the only way for the defense to effectively refute the testimony of a DNA expert is by presenting a competing DNA expert: “For the most part, the only attack I have really seen on DNA is if you have another expert come in and challenge the methods that were used – whether something is contaminated or the way it was drawn is improper or the vessels that held the analysis were somehow contaminated – that kind of thing. They actually just put it in the machine in the lab and let it run its course and these are the results and this is the profile and this is the match and that kind of stuff is pretty tough to refute.” The need to counter DNA testimony presented by the prosecution with an opposing expert defense witness was further emphasized by a handful of judges:

You have got to get a jury to question the process. If there are holes in the process, if they can’t trust the result, then they may not go with it. If they trust the result, it is done.

Another echoed a similar sentiment: “Sometimes the case really depends on it. The attorneys have no clue about what questions to ask and they really have a duty to have their own expert look at it.”

Interestingly, many judges did not speak about the substance of the testimony or the reliability of the evidence presented when asked to evaluate DNA evidence; instead, the analysis focused on the public speaking and presentation skills of the testifying expert. One said: “Some explain it better than others. Some just have a little bit more flair to their testimony because it is pretty dry stuff.” A second noted: “I haven’t seen one that I thought did such a bad job explaining it that it was discounted by the jury or should have been discounted by the jury.” A third articulated the importance of a clean and understandable presentation: “I think that if the expert is hyper technical it goes right over everyone’s head. The best experts are those that make it comprehensible to the ordinary juror.” Some judges complemented the presentation skills of DNA experts:

Most of them that I have seen have been from the Office of the Chief Medical Examiner. They have very clear English speaking witnesses that can articulate what this is, how it works and what it means . . . They’ve developed a script that is capable of explaining this very complicated stuff to the jury in a very simple way and defense attorneys don’t go into it because there is no reason for them to. They’re very competent in testifying.

While the ability of any witness to effectively convey their testimony to the trier of fact is of utmost importance, the major focus on presentation skills, as opposed to scientific processes, was odd; my questions explored the value of DNA evidence as an evidentiary tool, yet many of the responses were focused on witness showmanship instead of substantive testimony. Similarly, Brodsky et al. (2010) found that experts are often evaluated based on factors such as likeability as opposed to factors related to their expertise.

While the vast majority consider DNA evidence to be the “gold standard,” approximately one third of the judges agreed that DNA evidence is not capable of satisfying all of the elements of the crime claiming, instead, that it is merely one piece of the puzzle: “All it really shows is that someone with [the defendant’s] DNA was at a particular scene. As long as it’s for that purpose, I think it has generally been reliable.”

Another noted:

DNA doesn’t necessarily determine the case; it only determines defendant’s presence. What happened and everything else is still up for grabs. Sometimes you get DNA from a woman who says that she was raped; you still have to show that it’s not consensual. That’s not the end of the case.

A few stated that they have only seen DNA evidence effectively refuted on the grounds of contamination or chain of custody issues. For example, one argued:

You cannot challenge the science of DNA; everyone knows how reliable DNA is. I have seen cases where the evidence was mishandled or mixed with other fluids and then that type of issue comes up and then you get a mixed sample.

An analysis of judicial comments regarding DNA evidence reveals several recurrent themes. First, inculpatory DNA evidence increases judicial confidence in guilt (DNA provides “certainty” and “it’s always best to be sure”) more than any other piece of evidence, forensic or witness-based. Judges characterized DNA as “a rock crusher” and “100% reliable” (see Smith et al. 2011, where jurors rated DNA and fingerprints as most reliable). Yet, this certainty is juxtaposed against a lack of understanding of the science behind the evidence. In addition to implicit comments, almost a quarter of the judges explicitly concede that they do not understand DNA. Without knowledge of the science, many judges blindly accept the statistical probabilities espoused by experts without ever questioning the process. Second, while most judges welcome and are comforted by the perceived certainty of DNA evidence, they are critical of jurors for

expecting and wanting that same certainty (partially because it complicates case processing). In fact, criticizing the CSI effect is normative among judges. Finally, the vast majority perceive DNA to be extremely objective; yet, a majority of the discussions focused on subjective evaluations of the clarity and showmanship of the expert witness instead of a focus on the scientific principles involved. In fact, it is common for judges to accept the science of DNA at face value and direct their attention to their area of expertise -- evaluating showmanship and other witness presentation skills. Thus, while judges praise the objectivity of DNA, the majority of their perceptions are based on subjective evaluations of factors external to the science itself.

Judicial Perceptions of Fingerprint Testimony

Most judges view fingerprint evidence as a consistently reliable form of scientific evidence (see Smith et al., 2011). Ranked second to DNA by a majority of judges, fingerprint testimony was characterized by several as a ‘powerful’ tool at trial: “Fingerprint evidence, when presented by responsible witnesses who are experienced, and it is something where the jury can actually see how the prints match up and the number of points of comparison and all of that, is very strong evidence. I don’t think I have seen one where any serious questions were raised as to the accuracy of the evidence.” While highly regarded, most perceive fingerprint evidence to be more subjective (for a discussion on the subjectivity inherent in latent print analyses, see Dror, 2013) and weaker than DNA, though opinions vary somewhat. For example, one stated:

Fingerprint evidence is less reliable than DNA but second on the list. There is more subjectivity in terms of expertise. The analysis is not totally scientific. There are more personal opinions and it is considered less reliable. Experts in fingerprint cases are more tentative. They speak about a sufficient percentage.

Another said: “The DNA goes in the computer and it’s a match by the computer and it makes the match very accurately. With fingerprints, the computer says it is a match and a human being has to look with comparison microscopes and make a determination so there is more room for human error. “ In contrast, a few judges explained that they prefer fingerprint evidence to DNA since they can better understand it and visually observe the comparison in fingerprint cases. Further, one commented:

I find it more reliable than DNA evidence. It’s been around for a hundred years. We’ll see where DNA evidence is a hundred years from now.

Judges expressed that the strength of fingerprint evidence depends upon a variety of factors. First, several stated that there is a significant variation in the quality of fingerprint experts. While some discussed the scientific component of the fingerprint expert’s testimony, others were purely focused on the presentation skills of the witness:

It all depends on the quality of the expert, which varies. I really think that there’s a strong bearing on who collected it, how they collected it, what their background and training is and how they come across as a witness and how the person, who is now the fingerprint expert, which may be a different person, how they come across as a witness. Some people have better background and training than others and some people make a better witness and can explain things to a jury better than others. They’ll hold a chart up of this person’s actual fingerprint and then they blow up whatever and they will show you the lines of similarity and things like that. Juries can see for themselves; does it look like a match or doesn’t it, and you can kind of use your own common sense as well.

Many judges focused on the expertise of the witness. Conclusions regarding witness expertise varied:

I think it really depends on who has taken the prints. I have seen such slip shot police officers that should never have been police officers let alone technicians taking evidence and then trying to present it as matching or something. When they have only done 5 or 6 or 7, that doesn’t mean that they are an expert; that just means they happened to be there and somebody said, ‘here, you do this.’

Another noted: “There is skill in the laboratory and there is skill in testifying. I can only judge their skill in testifying. There are some people who can communicate more effectively than others, but I don’t know if that means they’re less skilled at achieving the right result.” About a third found the quality of the experts relatively uniform across the board: “The people from the police department do a good job in reading fingerprints and it’s usually uniform. I mean you don’t get a great variation in testimony.” Another said, “At the end of the day its just, I’m the expert, this is what it shows. Distinctions in the quality of these experts are minimal.” Yet, contrary to these judicial perceptions, scholars have found inconsistencies with latent fingerprint analyses among and within examiners over time (Dror, 2013).

A second factor determinative of the strength of fingerprint evidence is the number of points of comparison: “In some cases the fingerprint evidence was barely enough for a match and I thought that it was not persuasive evidence. At other times, it is a clear match and places the defendant at the scene.” Another said:

I don’t remember the number of points but I know they want a certain number of points if you are going to be certified; I don’t remember if its six or eight. I don’t recall off the top of my head but certainly a greater number of points would be much more favorable to the prosecution and conversely, less would be less damaging.

Another differentiated between partial and complete fingerprints:

Obviously, if you have got a partial print and you only have six points of analysis, instead of eight or twelve or whatever it may be, it may not be conclusive. A smeared fingerprint that’s only a partial is not going to be conclusive. If you have a really standard print with all of the ridges showing, I have never seen anything to show me that it’s not very reliable.

Finally, one judge pointed out that the fingerprint standards are not uniform: “You find out that some jurisdictions look for X number of points and others look for 2X.”

Similar to DNA, several judges expressed that fingerprint evidence is “only one piece of the puzzle”: “There is always the question of how did the print get there and that’s the piece that forensic evidence does not work on and that’s the job of the lawyers to figure out how the fingerprint got there.” Another emphasized the importance of fingerprint evidence when defendant’s identification as the perpetrator is at issue: “There are times when fingerprint evidence was so important because it’s a circumstantial case and it proves that a person is that person with this fingerprint which may be a clear match to your defendant and it puts him on the scene. At other times it has not been very useful at all.” Similarly, “depending on the case, just because you find somebody’s fingerprints, that in and of itself doesn’t necessarily mean anything unless it ties in with other evidence.” Finally, one judge compared fingerprints to DNA in explaining that probative value is case dependent: “It’s the same thing with DNA as it is with fingerprints. It’s like finding semen in a rape victim. Was it forcible or was it voluntary.”

In sum, fingerprint evidence is considered by most judges to be a reliable form of scientific evidence that can resolve identification issues. While viewed as more subjective and less reliable than DNA by most, some judges prefer fingerprint evidence since it is visually approachable. Judges appear to comprehend fingerprint analyses better than DNA; they can observe the actual prints, make their own visual comparisons and count the number of match points. This understanding allows for more critical analyses of the procedures utilized in a particular case. In contrast, DNA analyses are performed by a computer and undergo a process largely incomprehensible to most judges. Rather than

instilling doubt, this lack of understanding seems to result in a blind acceptance of the scientific processes involved. Finally, similar to DNA, and a consistent theme among all forms of forensic evidence, the majority of emphasis on fingerprint expert testimony was on subjective assessments of witness presentation skills and showmanship, as opposed to the substantive scientific processes involved.

Judicial Perceptions of Ballistics Evidence

Judicial opinions regarding the quality and reliability of ballistics evidence are somewhat mixed. Still viewed by most as a relatively reliable form of scientific evidence despite its recent debunking according to the National Research Council (Lichtblau, 2004), it still ranks below DNA and fingerprints in evidentiary strength by the majority of judges. While most find it reliable in terms of establishing the operability of a gun (i.e. whether a gun is capable of readily discharging ammunition), approximately a third of the judges expressed concerns regarding the ability of ballistic science to definitively determine that a particular shell casing was fired from a particular gun. Further, similar to other forms of forensic evidence, about a quarter of the judges cautioned that ballistic evidence is merely one piece of a case that is often not determinative of the ultimate outcome. Overall, judges were very complimentary of the ability of expert ballistics witnesses to testify effectively (as compared to fingerprint cases, where there was more variation in perceptions regarding presentation skills). Each of these points will be elaborated below.

First, most judges regard ballistics evidence to be probative and reliable. For example, one characterized ballistics evidence as “kosher,” while another described it as

“reliable and concrete and based on some physical reality as opposed to statistical analysis.” While a majority perceive ballistics testimony regarding gun operability (i.e. the ability of the firearm to discharge ammunition) to be reliable, some caution that this evidence is not very useful since operability is often not an issue. The reliability of expert ballistics testimony matching ammunition to weapons is also questionable in the minds of some of the judges. For example:

There are many more individuals in the world than there are guns and one is a product of creation and the other is a product of manufacture. I am not sure I accept that every barrel of a gun that is machine tooled is unique whereas I am more prone to believe that individuals are unique in that respect. I don't have any scientific basis for that; it's just a gut feeling.

In contrast, others stated that they believe that ballistics evidence is effective in matching ammunition to guns:

It's reliable evidence . . . My experience has been that microscopists and the ballistics labs are very good and for the most part they are able to say, if they have the gun and the bullets, that these bullets were fired by this gun or match the shell casings. It's reliable evidence.

One judge characterized ballistics evidence as a combination of stronger objective evidence and weaker subjective evidence: “In terms of gun operability, ballistics evidence is purely scientific and reliable. But if you are trying to match gyrations on shells, it is somewhat subjective.” This analysis is consistent with reports criticizing the ability of experts to accurately perform ballistics matches (see Lichtblau, 2004).

Perceptions regarding the value of ballistics evidence varied as well. Specifically, some judges cautioned that ballistics evidence has limited value. One judge stated: “I have never found ballistic evidence to determine guilt or innocence” while another noted that ballistics “is frequently just a small piece of the puzzle.” In contrast, others pointed

out that is can be helpful to the prosecution in establishing the case. As one judge described:

I think it's probative especially in shootings. You usually have a jury's attention. I think it is one piece of it but for the most part I think that people are listening. . . . Many times your witnesses are involved in the criminal world, as is the defendant. Then the ballistics comes in and it helps to kind of start to bring the story together a little bit.

Similar to DNA evidence, judges focused primarily on the expert's presentation skills instead of the substance of his or her testimony when evaluating this form of evidence. Most perceive ballistics experts to be skilled expert witnesses, with little variation in quality between witnesses. For example, one stated: "The DAs have a good handle on it because the same witnesses come in all of the time. It's almost like you press play and they say what they say." A second agreed and elaborated: "They testify very well because that's all that they do. So I think that their testimony is clean and succinct and persuasive." According to a third, "The ballistic experts are always trained together – trained to look at and connect with the jury."

In sum, most judges perceive ballistics evidence to be a solid form of forensic evidence that can assist the prosecution in establishing the defendant's guilt. While judges generally perceive ballistics to be reliable in establishing gun operability, they are less confident about the ability of a ballistics expert to achieve individualization (e.g., to match weapons with ammunition). Notably, matching requires human judgment calls and assessments whereas operability tests do not require these same determinations. Ballistics determinations are easier for judges to understand than DNA; with some basis for how examinations are performed, they are more critical of the process and somewhat less likely to take the evidence at face value. In fact, most tend to be more critical of forensic

evidence in general when they understand the mechanics and they tend to be more blindly accepting when they do not. Again, consistent with other forms of forensic evidence, judges focused on the testifying skills of these witnesses more so than on the content of their testimony. Evaluating presentation skills is an area of judicial expertise. Few distinguished the substantive value of this evidence from the showmanship, clarity and ability of the ballistic expert to relate to the jury. Thus, factors external to the substantive evidence itself appear to be driving judicial perceptions of the accuracy and strength of the scientific processes involved.

Judicial Perceptions Regarding Evidence of Chemists

In criminal prosecutions, chemists often provide expert testimony identifying case evidence as containing a particular type of drug (e.g. heroin, cocaine, etc.). Of all of the forensic forms of evidence discussed thus far (i.e. DNA, fingerprints, ballistics), chemist testimony is the subject of the greatest amount of criticism. Moreover, judicial opinions vary widely in this area with regard to reliability, with a significant number on both ends of the spectrum. As opposed to other evidentiary forms, they focused not only on the presentation skills of the experts but, also, on the reliability of the scientific procedures utilized and the manner in which they were performed.

The range of opinions regarding the reliability of chemist testimony is wider than the range in any other forensic evidence category. Approximately a quarter of the judges categorically stated that they find the testimony of chemists to be largely unreliable. As opposed to other forensic evidence categories, where judges were primarily focused on the presentation skills of the witness, criticisms regarding chemists additionally focused

on the scientific procedures utilized to reach the results. One stated: “We’ve had so many cases where the labs have been flawed. The chemists have been flawed. That is worrisome” (see Biben, 2011, for a report of flawed drug lab procedures in New York). Another spoke about inconsistencies between initial screening tests and subsequent lab examinations:

We take a lot of urine screens. One thing that is troubling in drug court is that we will get a positive screen when we take the urine screen and then we will send it to the lab for a more sophisticated test and it will come back negative. It happens too much to be right and it’s kind of troubling when that happens.

A different judge stated:

So many of the chemists are screwy. They have had issues in a lot of the labs. They haven’t met standards. Some of the chemists have fudged the numbers. It gets a little hairy. You have no way of knowing as you sit here that this is reliable or unreliable. I have seen them effectively impeached.

In contrast, some trust the testimony of chemists. For example, one judge stated that: “In all of my cases, the testimony of the chemist was always conceded by the defense . . . They concede so as not to distract. Stipulations on uncontested matters create good will.” Another stated that he finds the testimony of chemists to be more compelling than the testimony of other forensic witnesses: “The chemists are probably a little bit stronger because they just go through twenty minutes of how they clean their area and there is absolutely no possibility of contamination.” According to a third: “The chemist comes in and they talk about the presumptive test and then about the gas test, and how they confirm that it is cocaine or it’s heroin or whatever. That is very routine stuff and I have never seen anybody effectively challenge one of the police lab chemists on a simple drug analysis.”

More prevalent than in any other evidentiary category, judicial complaints regarding the communication skills of chemists are widespread. The most common criticism is that chemists often are weak in the English language: “As far as the testimony is concerned, sometimes they have trouble being understood because English isn’t their first language.” Another stated: “The biggest issue I have with chemists is that many of them cannot communicate well because of the language. There is an art to testifying that some people do not get, like the ability to hold the juror’s attention and being up on the field.” Another judge simply commented: “Chemists don’t make the best witnesses.”

Several judges expressed that there is a great deal of variation in the reliability of chemists. Numerous judges spoke about the contamination issues that were found in New York laboratories (see Biben, 2011) and blamed it on the lack of skill or sufficient training on the part of the testers. As one judge stated:

Some of these witnesses were police officers for a very long time. They weren’t private chemists, or doctorates – they were cops who received some training. They did the best they could, but if you are not properly trained in the field and that becomes your job, it depends on the day and it depends on the examiner.

As with all other forms of forensic evidence, some judges emphasized that the testimony of chemists is often not the primary evidentiary issue at play: “Drug cases usually rely on informants. Attacking the informant’s credibility and reliability and the adequacy of the police investigation is probably a stronger tactic than the chemistry.” Another stated: “Chemists usually come up in the drug cases. Usually the defendant stipulates that the substance is cocaine or whatever it is because that is really not a dispute and the issue is really their role in the transaction.”

Of all of the forensic evidence reviewed thus far, the testimony of chemists is the most heavily criticized by judges. Yet, there is a wide range of opinions. Some respect,

trust and rely on the process while others are concerned about the reliability and honesty of these experts. It is notable that judges spoke about the potential for substantive mistakes when speaking about chemists more so than they did in any other category of forensic expert. Yet, little reference was made to their own observations and experiences. Nor did most judges describe the processes they deemed unreliable or suspect. Instead, over a third discussed the publicity surrounding the police lab in New York (see Biben, 2011). It was the reputation of the evidence, more so than personal experiences with the evidence itself, that appeared to drive judicial perceptions of chemists.

By far, the most common criticism of chemists was their inability to speak English well. Given their strong focus on testifying abilities, these language skills irritated judges. The combination of poor presentation skills and a tarnished reputation due to crime lab errors relegated chemist testimony to the bottom of the forensic evidentiary pyramid.

Perceptions Regarding the Need for Forensic Evidence

A common impression among most judges is that juries over-rely on forensic evidence. Indeed, many were very critical of juries in this area and numerous judges complained that the CSI Effect hinders efficient and just case processing. Judges routinely answered questions posed regarding *their* own personal perceptions of the probative value and evidentiary necessity of forensic evidence by diverting the focus to the CSI Effect; this was consistent despite numerous requests on my part for them to discuss their own views regarding the evidentiary need for forensic evidence instead of jury views. When redirected to discuss their own perceptions, a majority of judges stated

that forensic evidence is helpful in establishing the defendant's guilt but that its importance varies by case characteristics and its use is not always possible, practical, or in some cases, necessary.

Some judges expressed the need for forensic evidence in certain types of cases. For example, one stated that "certain cases should have forensic evidence. Homicide cases for sure." A second commented that: "If there is no DNA and somebody is charged with rape or any sort of assault that's a big deal." A third noted:

It is always better to have forensic evidence. In sex cases, you need to match the sperm. That's pretty powerful evidence. DWI, if the driver refuses the breath test, it makes it much more difficult to get the conviction.

A fourth argued: "Obviously, if you have a body that is full of bullets you are going to think that the firearm analyst ought to be there checking them out and making sure they are from the same gun. If you have a case where the defendant has blood all over her shirt after a stabbing, you certainly expect for the DNA lab to confirm whether or not it is the victim's blood." Finally, a fifth stated: "Forensic evidence takes the burden off. It's so much easier to convict or not to convict someone depending on the forensic evidence."

Several judges stated that forensic evidence is not necessary in every case. For example, "you can have a case proven beyond a reasonable doubt without forensic evidence. I fought the forensic thing as a prosecutor for twenty years. I am not so awed by it." Another stated: "Forensic evidence is not always necessary. If there is witness testimony, and several witnesses saw whatever they saw, then I don't think that having a DNA hit off of a glass that was in the place is going to change anyone's mind that the three people that said they saw him really saw him."

Judges highlighted certain “valid” reasons that forensic evidence would be absent in a particular case. A widespread reason cited among judges was that “it is cost prohibitive to do it in every case.” Another explained that “on the lower level felonies or the misdemeanors the people and the police are not going to spend the money for certain testing as they would on a very serious case. It’s just not cost effective.” A second reason judges expressed for the lack of forensic evidence in certain cases is that scientific evidence is simply not obtainable in certain instances:

I tend to rely less on the absence of forensic evidence than others might. If there is no fingerprint evidence, there could be 1001 reasons why. You can still have a good case.

In sum, most judges praise forensic evidence and the probative value and certainty that it brings. While the perceived need for it varies among crime type (i.e. judges expect it more in homicide and rape cases), there is a strong acceptance and understanding that it is not always practical, possible, or cost effective. The reliance by jurors on forensics was highly criticized by judges and it seemed almost politically correct for some judges to distance themselves from “television watching jurors” who unreasonably hamper case processing by demanding forensics. This tendency to downplay the necessity of forensics and to criticize those who demand it stands in sharp contrast to expressed sentiments regarding the strong evidentiary value and confidence in guilt that science brings to the courtroom.

B. Judicial Perceptions of Witness Testimony

Judicial perceptions regarding the probative value, credibility, reliability and general strength of witness testimony vary based on the type of witness testimony

presented and case characteristics, among other factors. In this section, I will explore judicial perceptions regarding witness-based testimony. First, I will distinguish four different types of witnesses: eyewitnesses, character witnesses, alibi witnesses, and psychologists. Following this analysis, I will explore the factors that judges consider in reaching credibility and reliability determinations with regard to these witnesses.

Judicial Perceptions of Eyewitnesses

Judicial perceptions regarding the value, credibility and reliability of eyewitnesses are mixed. While judges value the “human element” inherent in witness testimony and the majority of judges find eyewitnesses to be credible or truthful most of the time, consistent with the scholarly literature, most judges have deep concerns regarding the reliability or accuracy of eyewitness testimony (Loftus, 1974, Sander, 1984; Sharps et al., 2007; Shermer, Rose & Hoffman, 2011). While a minority of judges spoke about their own experiences regarding eyewitness reliability, most initially focused on published studies or media reports and were subsequently redirected to discuss their own experiences. Some judges emphasized the need for corroborating evidence, particularly in one-witness identification cases. Finally, several explained that the accuracy of eyewitness identification testimony depends on the specific details and varies from case to case. These perceptions are discussed below.

First, many judges characterized eyewitness testimony to be a “scary,” “weak” or “frail” form of evidence. For example: “Eyewitness testimony has to be taken with a grain of salt. The fact that people who are in a stressful situation observe things

differently and testify differently as time goes on . . . there are a lot of factors at play where you can't say, 'this is absolute fact.'" A second stated:

Eyewitness testimony can be just horrific. I mean five people watch an incident and you get complete opposites. You just think how can they all be in the same room and give completely different versions? You always have to scrutinize eyewitness testimony.

Many judges focused on the studies regarding eyewitness identifications when discussing their perceptions of eyewitnesses and appeared to be highly influenced by media reports and academic literature (see Loftus, 1974, Shermer, Rose & Hoffman, 2011). For example, one stated: "It's very difficult. You read the literature that talks about misidentifications and then you listen to a witness and they are absolutely convinced that their identification is correct and you know that you have recited a thousand times that accuracy is not the same thing as confidence and you just know that they are not interchangeable but you know that the studies are troubling." In contrast, several noted that eyewitness testimony can be "very effective" and is often "a necessary component of a trial." A handful spoke about the importance of corroboration in cases involving eyewitness identification. For example, one claimed: "I don't put 100% credibility in identification testimony. I do put a lot of credibility in identification testimony when it is corroborated by other factors . . . DNA, several eyewitnesses, [defendant] caught quickly in the area." Another judge stated: "As the years have gone by people realize that it is only good if it is corroborated." Finally, a third stated: "Most cases need corroboration. Personally I want more. It is more of a reliability issue -they may think that they are sure - than a credibility issue."

Judges discussed the challenges involved in one-witness identification cases. In discussing these cases, most noted that, from a legal standpoint, the identification

testimony of a single eyewitness is sufficient to establish the defendant's guilt beyond a reasonable doubt. Despite its legal sufficiency, a majority stated that one-witness identification cases are troublesome and difficult cases to prove. For example:

I think one witness IDs are built in reasonable doubt. Unless it's somebody that you know. But I think your usual quick criminal encounter is really hard. I find them the scariest cases.

A few mentioned that one-witness identification cases often result in acquittals. For instance: "Usually they don't get convictions unless the defense goes ahead and does something stupid." In contrast, a few focused on the legal adequacy of one-witness identifications:

I don't need corroboration of a witness's testimony because the law tells me that it is not required. Do I think that it would be beneficial? In many cases, absolutely. If I'm asked to do my job without it, could I? Sure.

Judges stated that the reliability of eyewitness identifications often depend on the specific factual circumstances of the case. Some of the factors articulated include the length of time the eyewitness observed the defendant, whether the witness was able to give a detailed description, and the lapse in time between the criminal event and the identification (see Loftus, 1974; Sander, 1984; Shermer, Rose & Hoffman, 2011). For example, for one judge, corroboration is generally necessary "unless it's one witness who was kidnapped for four days and had plenty of time to look at the person."

In sum, judicial perceptions of eyewitness testimony seem to mirror public perceptions and media accounts. In contrast to forensic evidence, which is often presumed to be reliable absent evidence to the contrary, preconceived notions regarding the potential for human error inherent in eyewitness testimony renders it suspect from the start. While the value of the human voice resonates with many judges, the lack of

certainty in these identifications stands in sharp contrast to the perceived certainty associated with many forensic evidentiary forms. Several judges explicitly concluded that uncorroborated one-witness identifications are always suspect; judicial confidence in these cases is the lowest.

Judicial Perceptions of Character Witnesses

In general, most judges perceive character evidence to be a weak, unhelpful form of evidence. For example, one judge stated,

Good people can commit crimes. They can still have good character. It doesn't mean that they didn't make the mistake and do what they did. Character evidence is usually not that helpful.

Another stated: "If you had to make a chart from the best evidence to the worst evidence, DNA would be Number 1 and character evidence would be at the very bottom." A third explained his reasons for concluding that character witnesses are irrelevant: "Even John Gotti had a mother who loved him . . . but I give more thought to what I had for lunch than John Gotti gave to killing someone but there were those people who loved him. So, you know, different people show different sides of themselves in different situations."

Some argued that the testimony of character witnesses must be carefully scrutinized because it is naturally biased. As one noted: "I think you have to look at them with a grain of salt because they are there to help, whether it's the people's witness or the defense witness. You know that their testimony is going to be slanted toward that perspective, but I would say that most of these people are honest people who come forward and give you what their opinion is."

A minority of judges stated that they believe that character witnesses can provide valuable evidence, but many of these judges emphasized that the value of this evidence is largely dependent on the characteristics of the witness: “When the pope is a character witness, that’s impactful.”

A consistent theme was that the evidentiary rules in New York, with regard to character witnesses, limits its use significantly (see N.Y. Criminal Procedure Law Section 60, 2015). For example, the permissible focus of character witness testimony is on the witness’s reputation for truthfulness in the community, which has limited value. Further, it is often strategically dangerous for the defense to put a character witness on the stand. For example, one judge stated that,

character witnesses have absolutely no value. It is stupid because of the limitations of what a character witness can testify to. All he can say is that he heard that the defendant has a reputation for truthfulness. Then the other side can ask about bad acts and the defendant’s record and that destroys the character witness. The dangers of putting a character witness on the stand are high and the potential benefits are small.

While analyses of judicial comments related to character witnesses reveal that little weight is placed on this form of testimony, several judges explained that high status character witnesses may be influential evidence (e.g. the Pope). Thus, defendants capable of attracting these high status witnesses are better able to present a strong defense. Cooney (1994) argues that evidence is socially produced; that is, evidence in criminal cases is dependent on law enforcement priorities and efforts, as well as the ability of victims and defendants to attract witnesses in support of their cause. Thus, the status characteristics of the parties to the offense actually *produce* specific types of evidentiary packages (Cooney, 1994). Consequently, defendants able to secure high status character witnesses are in a better position to defend their case.

Judicial Perceptions of Alibi Witnesses

More so than any other witness based evidentiary category, judges gauge the value of alibi witnesses on external factors surrounding the substance of the alibi. These factors include the nature of the witness's relationship with the defendant, the timing that the alibi information was brought to the court and prosecutor's attention, and whether or not there is documentary evidence available to support the alibi. Moreover, judges perceive the quality of alibis to vary dramatically. As one judge stated: "The right alibi can be a very powerful defense while the wrong alibi can effectively shift the burden of proof from the prosecution to the defense." These concepts are discussed below.

First, the nature of the relationship between the defendant and the alibi witness was the key focus in evaluating the quality of this evidence. For example: "Most alibi witnesses have an interest in the case because they are a friend, or mother or other relative. They start off behind the eight ball because they have an interest because they are connected to the defendant." Another judge stated: "I've seen cases where the mother comes in and says, 'he was home watching television with me. A mother wants to protect her kid so it doesn't hold much weight'" (see Culhane & Hosch, 2004, which found that mock jurors are more likely to acquit if the alibi witness does not have a relationship with the defendant). In contrast, a third judge understandingly noted that it logically follows that many alibi witnesses have relationships with the defendants: "After all, who are you with most of the time?"

Judges categorize certain alibis as strong or "ironclad." Examples provided of successful alibis include proof that the defendant was incarcerated at the time that the crime occurred, proof that the defendant was on a flight, and testimony that the defendant

was at work during the commission of the crime. Alibi witness testimony in these examples can be supplemented with documentary evidence such as plane tickets, prison records, and work time cards.

Judges noted that the timing that the alibi is brought to the court's attention is critical: "I think people see through alibis. When they come in the eleventh hour and they have limited details, people see through them." Others stated that credible alibi evidence presented early in the process could result in the case being dismissed by the prosecutor.

Judges further emphasized that alibi witness often are impeached during cross-examination and impeached alibi witnesses could be very harmful to the defense.

Accordingly:

It's gotta be rock solid. The slightest deviation will destroy it. It's all or nothing. I have had ineffective alibis in my courtroom. My view is that it has the potential to do more damage than anything the DA offers. If a jury thinks that you put on a bullshit alibi, they will not forgive you.

In sum, judicial perceptions of alibi witnesses varied according to the nature of the alibi. Consistent with other evidentiary themes observed, alibi witnesses whose testimony is corroborated by other evidence (e.g. plane tickets, records that defendant was in jail at the time) are more likely to be believed than alibi evidence based on witness testimony alone. Further, credibility considerations, such as motive, are likewise considered by judges in assessing the value of these witnesses. Viewed as a subjective evidentiary form, in which the motivation to lie to protect the defendant is high, uncorroborated alibi witnesses are perceived as weak evidence.

Judicial Perceptions of Psychological Testimony

While judges vary in their views regarding the strength, need for, and probative value of psychological testimony, the majority of judges expressed a low level of confidence in the value of psychologists in criminal proceedings (see Faust & Ziskin, 1988, for a discussion of the limited value of psychological testimony in the courtroom). In fact, even among judges who perceive that psychological testimony is helpful, the vast majority qualified and limited their opinions to specific situations. Characterized by approximately a quarter of the judges as a “soft science” where you could “say anything,” participants expressed a host of reasons why they place little value in this form of testimony.

First, more than half of the judges characterized psychological testimony as “subjective.” As one stated:

That is really subjective. You don't even have a smudge to look at like you do in a fingerprint case. Even well qualified people are simply giving their own subjective views.

Other judges focused on the fact that psychologists are usually paid to testify by either the prosecution or the defense and are biased toward their employer. For example: “If you are getting paid by somebody and you want to have a livelihood, you are going to have to testify for their side.” Another judge categorically stated that “experts hired by one side are biased and lack credibility.” Others focused on how “dueling experts” confuse matters;

Well, I think their value is definitely one that helps the defendant or helps the prosecution; depending on what side they are on. You get the testimony that you are looking to get . . . That is a very subjective profession and it's something that, if you want a defense slant you can get it and the prosecution has their psychiatrist that they deal with on a day in day out basis so that they can get the testimony they want and the jury is left confused.

Another focused on the inequities of paid experts: “The poor have it covered through Legal Aid and the rich can afford it. But the middle class can barely pay their attorney fees, let alone expert testimony.” Interestingly, the negative implications of “hired experts” were a greater focus when discussing psychologists than for any other scientific expert. This difference appeared to be fueled by overall skepticism about the field of psychology.

Some judges expressed opinions regarding the limitations of the field of psychology in general:

It’s all hocus-pocus. My God, I mean, how do these people even have a job . . . I don’t put a lot of stock in it . . . There are good individuals that will give you broad strokes but analysis or opinions or definite conclusions as to why somebody has done something, they have no clue.

Yet another, who indicated his preference for the “harder” sciences, stated: “I find psychiatrists a little more credible than psychologists because at least they have a real medical degree.”

A major criticism expressed by several judges is the constantly changing nature of the field. As one stated:

There are no definitives in psychiatry. Having been a witness to DSM I, II, III, IV and now I think we’re up to V. I won’t say it changes with the wind but it changes.

A second expressed a general lack of understanding on the part of psychologists as to what is transpiring. He stated, “It seems to me that these psychiatrists are not really sure about what is going on and it keeps changing. It is almost trial by error . . . I am a little bit more skeptical now than I was years ago.” Finally, a third combined his generally negative views of psychology with the challenge of “dueling experts.” For example, he

claimed: “It’s a popularity contest with competing experts because there is no real science behind it. The most likeable psychologist wins when you have dueling testimonies.”

A minority of judges expressed their views that psychology is a probative and reliable form of evidence. Yet, virtually all of the judges who expressed generally positive sentiments qualified their positions. For example, some focused on the quality of the particular expert:

It can be very helpful. Depends on the quality of the individual, the quality of the perception. It’s an area where you get lots of disagreement. It’s a lot harder to pinpoint. As soon as you start talking about the mind, and the operations of the mind, you’re not talking about things as evidently objective as the DNA or fingerprints or ballistics. It doesn’t mean that it’s not potentially very useful but it really depends much more on the quality and insight of the individual giving the opinion. It can be very useful but trickier.

Another stated that while psychology has value, psychological testimony is generally presented with a slant toward the represented party: “They offer valuable testimony but you have got to sift through it to see what the actual text summary is, as opposed to the slant that they’re trying to put on it.”

Some judges delineated certain areas in which psychological testimony is helpful to the trier of fact. Specifically, psychology can be helpful in making diagnoses, but is limited in its ability to predict future behavior:

I still think psychology is not as much of a science as everybody wants you to believe. I think there are some certainly serious mental health diagnoses that can be made and that people can be fit into categories. I think that predictability is something that none of them can do accurately unless it’s the more violent somebody is, the more likely he is for violence. Other than that, I really don’t see it.

The nature of the case influenced the value of psychological testimony for a handful of judges. For example, several focused on the benefits of psychological testimony in

trauma cases: “I think it can be helpful depending on the trauma. If a witness has an inability to even verbalize or articulate what exactly happened, many times there can be a reason for it . . . it’s a piece of it.” Another judge described an incident regarding rape trauma and expressed both the usefulness and the potential pitfalls of psychological testimony in this area: “I think that it’s very useful evidence for the prosecution and I think that jurors tend to credit that type of testimony. The problem that I have with it is to avoid it simply bolstering the testimony of the victim or the complainant in the case because it’s a very fine line . . . I find it important evidence when it’s relevant, I just don’t find it relevant in every situation.”

Similar to other expert testimony, the testifying skill of the particular psychologist witness was a key factor in the value accorded to the testimony for certain judges: “ It depends on the witness . . . I could be talking gibberish but if I sound really good doing it I may project better than somebody who is talking real science who sounds like they are an idiot. That’s part of the whole aspect of trials in any part, is to the degree to which a trial is theatre, which it is, the performances of the actors make a difference.”

In sum, the majority of judges expressed skepticism toward the field of psychology in general and the value of paid psychologists in particular. The focus of judicial comments was directed to the inherent subjectivity of psychological testimony. Yet, similar subjective factors were present in some forensic evidence, even though judges viewed those evidentiary forms as more objective and reliable. First, judges complained that psychological testimony is based on subjective opinions; yet, fingerprint and ballistics comparisons necessitate a certain degree of subjective assessments as well and they are not subject to nearly as much criticism. Second, judges pointed out that

psychologists testify based on “the interests of the party who hire them.” Yet, dueling forensics experts are in the same predicament but judges did not express similar sentiments when discussing those forensic evidentiary forms. While forensic science is difficult for some judges to comprehend, psychological testimony is more approachable and, consequently, more open to criticism. Finally, similar to scientific expert testimony, judges focused on the portion of the testimony that was most familiar to them – witness presentation skills and showmanship.

Credibility and Reliability Factors

In determining the value attributed to witness-based testimony, credibility and reliability factors are key determinants. Different judges focus on various criteria in reaching these evaluations. Below I discuss a range of credibility and reliability factors considered by judges.

Credibility Factors

In assessing the probative value of witness testimony, the trier of fact must first determine whether or not the witness is credible. While some judges confidently explained the manner in which they determine witness credibility, others were tentative in describing both the factors that they consider and their ability to accurately assess credibility. For example: “It’s a very difficult area, it’s a very subjective area and I don’t claim to have any superior ability to determine credibility in the close case. And I have been wrong in the past.” Consistent with scholarly literature, judges discussed a myriad of factors that they consider. While consistency in testimony, motive, demeanor, body

language, attitude and witness background were the most commonly stated factors, judges vary widely in the factors that they deem important to credibility assessments (see Brodsky, Griffin & Cramer, 2010; Connolly & Gordon, 2011; Porter & Brinke, 2009; Wessel et al., 2006). Further, several judges emphasized the importance of considering the totality of the circumstances in reaching credibility determinations. The credibility factors described by judges are highlighted below.

First, more than half of the judges stated that the consistency of the witness is a key factor in determining witness credibility. Judges spoke about consistency in different ways. Some focused on the witness's ability to remain consistent during cross-examination and a few spoke about whether the testimony was consistent with common sense and other case evidence. However, others focused on the consistency between statements made by the witness at different case processing stages: "I would probably ... [look at] consistency between the statement given at the time of their first statement to the police, grand jury testimony, trial testimony, hearing testimony, and then how consistent they are from story to story."

Some judges focus on the witness's body language and eye contact in determining credibility:

I look at body language and I look at how they are moving their eyes. Which way are their eyes going as they are thinking. There is a whole science behind left bending eyes or right bending eyes; which is recall and which is thinking up a lie.

A central area of focus for judges in determining credibility is motive. One judge stated:

I look at the witness's motive to be involved. For example, a drug informant who got caught trying to make a buy from someone – they may be testifying in order to get a better disposition. I think that they can be inherently suspect.

Judges also consider the relationship between the defendant and the witness in determining whether the witness is motivated to lie to either protect or harm the defendant.

The witness's educational background and ability to comprehend the questions was also relevant: "It is important to evaluate a person's testimony in conjunction with their educational background; you may think that someone is being evasive or is lying when in fact they really don't comprehend. Then you need to slow it down and break it down and not be disrespectful toward them but just simplify your lawyer talk so that they understand you a little bit better."

About a third of the judges spoke about the witness's general demeanor in assessing credibility. For example:

As a rule of thumb if somebody is too calm I feel they were probably rehearsed and maybe they weren't telling the truth. If they were nervous they probably had a good reason to be. Not because they were lying . . . because you should be uncomfortable. I was once called as a witness and I was uncomfortable.

A few judges focused on whether the witness was evasive or forthright. One stated:

You know right away once people start beating around the bush. Like a parent asking a kid, "did you break the lamp?" The kid says, "well we were working in the living room" instead of just saying, "yeah, I did." When eyewitnesses start being vague, instead of just answering the question, when they beat around the bush like politicians do, that's a red flag to me.

In explaining the importance of forthrightness, another judge stated: "If they treat the prosecutor much differently than the defense attorney, it is a red flag to me."

A few judges spoke about the challenges of police witnesses and distinguished the credibility of police officers from civilian witnesses:

I think that civilians are usually credible unless there is some sort of agenda. When cases are lost it is usually because of a cop. Not because the cop is bad or lying but because they so want to put their point across that they will exaggerate. Then the jury sees that and the whole thing unravels.

The witness's criminal history as well as other background information is a factor articulated by several judges. For example, one judge noted: "Needless to say, you don't get your players from central casting." Another said:

The problem is that you are stuck with whoever is there. The drug dealers and the murderers and the rapists don't hang around with doctors and lawyers and people with PhDs who are honest, law-abiding citizens. They hang around with similar types of people in their environment.

Finally, a few spoke of the importance of evaluating the testimony in light of common sense: "If it's inconsistent with common sense and normal experiences in life, then there's a concern with respect to the credibility of that testimony."

An analysis of these credibility factors reveal that evaluations are based on subjective determinants that vary from fact finder to fact finder. For example, body language, eye contact and demeanor may be perceived in certain ways by one fact finder and in different ways by another. Even factors such as motive and the relationship between the parties may be interpreted in distinct ways. This unavoidable subjectivity is at the core of credibility determinations. For many judges, subjectively appears to be synonymous with weakness and subjective evidence is associated with uncertainty and doubt.

Reliability Factors

While a witness's credibility is an essential and preliminary determination that must be made in assessing the weight to afford to his or her testimony, the reliability of

the witness's testimony is a critical and distinct evaluation that must be reached by the trier of fact. Indeed, particularly in the area of eyewitness identification, history has revealed that even truthful and credible witnesses can be mistaken (Loftus, 1974; Sanders, 1984; Sharps et. al, 2007; Shermer, Rose & Hoffman, 2011). At least half of the judges discussed the existing literature regarding documented weaknesses inherent in eyewitness testimony when asked about their perception of the reliability of eyewitnesses. Similar to judge's discussions of the jury's vulnerability to CSI effects, I needed to redirect judges to talk about their own observations and perceptions, as opposed to the material they have read, on numerous occasions.

At least half of the judges indicated that assessing the reliability of witnesses – particularly eyewitnesses – was an extremely difficult task with an uncertain outcome. Particularly in the situation where the defendant and the identifying witness are strangers, correct identifications are difficult: “We want to be able to say this is the person that did it. It's hard. It's happened to me and there was no way that I could identify someone.”

Another stated:

I have to tell you that when I am trying a case and there's a jury I am happy. Let the jury sort it all out. It's really hard to say beyond a reasonable doubt that that is the person.

Some judges spoke about the importance of having other corroborating evidence to supplement eyewitness testimony, given the inherent uncertainty: “I look for verification with DNA and other physical evidence.”

Judges consider a host of factors in reaching reliability assessments, particularly when identification of the defendant is at issue. Some focus on the circumstances under which the identification was made, such as the length of time in which the observation

was made, “whether there a gun pointing in [the witness’s] face” or “whether they were looking intently and what was their reason for looking intently” whereas others focus on the details that the witness is able to provide. For example, one judge stated: “What they remember from what they saw, whether there was a certain thing that stood out on the person, or the clothing they were wearing; or the situation they were in; where they were, what was around them, what could they identify, what could they remember, how specific and clear could they be.” Alternatively, “were they able to give a fairly accurate description of the skin color and any distinguishing feature? I have had cases where someone clearly had a gold tooth or clearly had a huge scar on his or her face.”

Judges further espoused a myriad of other relevant reliability factors. These factors included the description given of the perpetrator at the time of the incident compared to the defendant’s appearance at trial, the witness’s ability to observe (i.e. lighting conditions, distance between the witness and the defendant at the time of the observation), level of fear at the time of the incident and whether the witness was under the influence of drugs or alcohol at the time that the observations were made. A few judges discussed the importance that the identification was made without any suggestiveness on the part of the police.

In sum, reliability concerns are paramount for judges in eyewitness identification cases. While reliability factors are helpful in guiding judges to reach factual determinations, the combination of judges’ personal experiences and exposure to the media render eyewitness testimony extremely fragile for most. Similar to credibility determinations, even carefully reasoned and explored reliability decisions are based on subjective evaluations of facts. Judges are painfully cognizant of the possibility that,

despite careful evaluation of credibility and reliability factors, their conclusions regarding witness testimony may be mistaken.

C. Judicial Comparisons of Evidentiary Strength Between Forensic and Witness-Based Evidence

The vast majority of judges perceive forensic or scientific evidence to be a stronger, more objective form of evidence than non-scientific, witness-based evidence. In fact, when asked to compare the strengths of these evidentiary forms, a majority unequivocally stated, without qualifying their comments, that forensic evidence is far more objective and stronger than other witness testimony. As one stated: “As a judge, forensic evidence makes my job easier. It prevents guessing as to whether the eyewitness’s subjective view is right.” Another claimed:

Forensic evidence is the ultimate. You don’t need anything else. I’ve had a conviction of a rapist without the witness having ever seen him, years later. Jury said, “DNA is his” and boom. Done. Another case of burglary. Someone breaks in through the kitchen window. No other evidence there. Can’t be identified by the complainant; but as he was exiting the window just a single drop of blood was left on the windowsill. Boom – got him. It’s the ultimate.

A third judge stated: “Science doesn’t lie. People make mistakes. People lie. People can deceive. The science doesn’t.” Finally, a fourth judge noted: “Well, scientific evidence, most of it, either is or it isn’t whereas witness testimony, no matter what they say, there is a possibility that they are wrong, right, half right, half wrong.”

Some judges distinguish between different forms of forensic evidence in assessing evidentiary strength. For example, one stated: “Not all forensics are the same. If you are talking about DNA or fingerprints then definitely, forensics is stronger. Chemists . . . not as much.” A few judges expressed their concerns that forensic evidence is still based on

expert witnesses and consequently, is subject to similar reliability and credibility concerns as lay witness testimony. In particular:

They are still just experts. They are individuals giving their opinion. I think the reliability of any given expert's opinion has to be tested, and it's not automatically persuasive. It's not necessarily more reliable than anyone else's testimony, certainly in my judgment.

Despite these concerns, the vast majority of judges perceive forensic evidence to be a more objective, stronger evidentiary form than non-scientific witness testimony.

Even still, several judges highlighted the benefits of non-scientific witness testimony as compared to forensic evidence. First, judges expressed the value of witness evidence that has the "human element." For example: "The human element gives a story strength." Or "there is a human element where the jury can identify with the witness."

For a third judge:

The human being is a brilliant thing. You see things, you hear things. You're perceptive. In fact, people are convincing in that regard. It's called the advertising industry.

Finally, "you certainly get a little more compassion and a little more concern and caring from a witness rather than something cold and black and white. But as far as it being superior to something black and white, that I don't see."

According to many judges, while credible and reliable witness testimony is sufficient to convict, it is always better to have corroborating forensic evidence. For example: "Forensic evidence is a complement to witness testimony. You need them both." And "the cumulative effect of a combination of different types of evidence is helpful."

In sum, while eyewitness testimony is associated with subjectivity, uncertainty and doubt, viewed as a whole, forensic evidence instills confidence in guilt by bringing perceived objectivity and certainty to the process. While there is some variation within each broad evidentiary type (i.e. forensic and witness-based), this theme is consistent throughout. Even judges who perceive forensic evidence to be unnecessary in certain situations still prefer to have scientific evidence included in the evidentiary package due to its perceived strength and objectivity.

D. Conclusion

A review of the qualitative data gathered during the course of these judicial interviews reveals that, as a group, New York State judges perceive forensic evidence to be a stronger and more objective form of evidence than witness-based testimony. This research further suggests that within the category of forensic evidence, judicial confidence in specific evidentiary forms is influenced, at least in part, by concerns about human error and the subjective nature of human perceptions, with judicial confidence increasing as the potential for human error decreases. These findings support the premise that judicial confidence in guilt is increased by the presence of perceived objective forensic evidence, while the subjective nature of other witness-based evidence results in uncertainty and doubt.

Judicial perceptions regarding the strengths of various types of forensic evidence reviewed in this chapter (i.e. DNA, fingerprint, ballistics, chemists) vary significantly. Observed differences often focus on the level of *perceived* objectivity associated with the evidentiary form. For example, DNA evidence is perceived as the strongest and most

objective form of forensic evidence. While the statistical probability of a match being accurate is extremely compelling in DNA cases and accounts, in part, for its high ranking on the evidentiary scale, much of the focus among judges was on the objective nature of DNA evidence. Specifically, some judges referenced the fact that DNA evidence is largely processed by machines with little human involvement and is consequently more objective than other evidentiary forms. For instance, “[t]he DNA goes in the computer and it’s a match by the computer, it makes the match very accurately.” Notably, several judges, who praised DNA for its objective quality, conceded that they lacked an understanding of the scientific processes involved. Yet, despite this admitted absence of knowledge regarding DNA analyses, these judges concluded that DNA is the strongest form of forensic evidence due to its perceived objective nature.

In contrast, computer identified fingerprint matches must be examined and evaluated by a fingerprint expert and therefore, are more subjective in nature. As one judge explained: “With fingerprints, the computer says it is a match and a human being has to look with comparison microscopes and make a determination so there is more room for human error.” The potential for human error inherent in fingerprint analyses is thus, arguably, higher than the possibility of human error in DNA cases. These findings are suggestive of the possibility that judicial perceptions of evidentiary strength are, at least partially, driven by concerns about human error and the subjective nature of human perceptions with judicial confidence increasing as human involvement decreases. Notably, despite articulating confidence in forensics due to its “objective” qualities, judges tend to evaluate forensic experts using subjective assessments unrelated to the scientific processes involved. These results are similar to Brodsky et al. (2010), who

found that expert witnesses are often evaluated by factors unrelated to their expertise. Some of the factors include believability, likeability and trustworthiness. Similarly, these judges discussed showmanship and courtroom skills in evaluating forensic experts. While praising forensic evidence for its “objective” qualities, the majority of analyses were based on non-scientific subjective criteria. Approximately a quarter of the judges conceded that they did not understand some of the scientific processes involved. Without this basis, judges simply accept the science and focus on the courtroom skills that are within their area of expertise. This is evident from the fact that forensic evidence that is somewhat more easily understood (fingerprint analyses compared to DNA) was criticized more heavily since judges had a basis for evaluating the evidence.

Consistent with these concepts, most judges rank ballistics evidence below fingerprint evidence. As opposed to fingerprint evidence, in which a computer first identifies a match and an expert later visually examines the suspected match, matching gyrations on bullets and shell casings is generally performed exclusively by ballistics experts and therefore, are more susceptible to human error. Yet, compared to the testimony of eyewitnesses, which has been shown to be unreliable at times and has resulted in wrongful convictions, forensic evidentiary forms are perceived by judges as far more objective. For example, “[scientific evidence] is more objective than a witness testifying who is subject to human frailties. Human error. That kind of thing. Whereas a nine millimeter bullet can’t lie. Won’t lie. A blood type, it is what it is. That kind of thing...” While chain of custody and contamination issues are concerns in forensics, they are far lower than judicial concerns regarding the “scary” and “dangerous” nature of eyewitness identifications. Consequently, most judges view witness testimony to be a

weaker, more subjective and less reliable form of evidence than forensic evidence. These findings support the premise that forensic evidence increases judicial confidence in guilt while witness-based evidence often creates uncertainty.

CHAPTER 6: THE INFLUENCE OF EVIDENTIARY WEIGHT AT SENTENCING POST-TRIAL

Judges may consider many legally relevant factors at sentencing post jury-verdict; however, they may not usurp the role of the jury by revisiting the issue of guilt (6th Amend., U.S. Constitution).⁴⁷ Indeed, the Supreme Court has reinforced the importance of preserving a defendant's Sixth Amendment right to have a jury determine whether all of the elements of the crime have been proven beyond a reasonable doubt (*Alleyne v. U.S.*, 2013; *Apprendi v. New Jersey*, 2000; *U.S. v. Gaudin*, 1995).

When a defendant exercises his or her right to a jury trial and is subsequently convicted, the sentencing judge is tasked with determining sentence based on a verdict that he or she did not render. In most cases, this structure is not problematic; judges often agree with jury verdicts (Kalven v. Zeisel, 1966; see also Eisenberg, Hannaford-Agor, Hans, Munsterman, Schwab & Wells, 2005) and are comfortable proceeding to the sentencing stage of case processing. Yet judges and juries do not always agree (Kalven v. Zeisel, 1966; see also Eisenberg, Hannaford-Agor, Hans, Munsterman, Schwab & Wells, 2005). While the vast majority of judges respect the importance of upholding legal jury verdicts, irrespective of whether they would have made the same decision,⁴⁸ the distinct roles of juries and judges in criminal prosecutions may place a judge in the challenging predicament of imposing sentence based on a verdict the judge believes to be incomplete

⁴⁷ A judge may set aside the verdict if there is no reasonable view of the evidence, viewed in a light most favorable to the prosecution, to support a guilty verdict (*Jackson v. Virginia*, 1979). This is a narrow exception that the vast majority of judges reported rarely (if ever) using. In such instances, the case would not even proceed to the sentencing phase.

⁴⁸ Judges in this study indicated that they have great respect for the jury's role. Even judges, with decades of experience on the bench, rarely or never set aside a jury verdict.

or incorrect.⁴⁹ This situation may give rise to a sharp internal conflict in circumstances where a judge lacks confidence in the defendant's guilt. For example, seriousness of the offense is an important and legally relevant sentencing factor (von Hirsch, 1992; Bushway & Piehl, 2001; Kramer & Ulmer, 1996); yet, it presupposes that the defendant actually committed the offense. Said differently, confidence in guilt is a prerequisite for valid sentencing concerns; only a guilty defendant poses a threat to the community and imposing punishment for the purpose of deterring crime for a specific offender necessitates the perception that a crime was actually committed (Paternoster, 2010).

This dissertation seeks to illuminate whether, to what extent, and how perceptions of evidentiary weight influence judicial confidence in guilt and ultimately impact sentencing decisions after conviction at trial.⁵⁰ Thus far, the analytic chapters (Chapters 4 and 5) have provided key findings that begin to answer these questions. In Chapter 4, quantitative analyses of violent felony trial convictions reveal that the type and quantity of evidence in a case significantly influences sentence length. Specifically, cases with forensic evidence and cases containing more pieces of physical evidence are associated with longer prison sentences for convicted defendants. These findings suggest that judicial perceptions of evidentiary strength – presumably operating through judicial

⁴⁹ This generally includes instances where a judge feels that a jury under- or over-convicted and only rarely involves cases where the judge believes that the defendant should have been acquitted.

⁵⁰ This chapter examines the influence of evidentiary weight (and resulting judicial confidence levels in a defendant's guilt) on sentence severity in trial cases through qualitative interviews with sentencing judges. The chapter focuses exclusively on trial cases for several reasons. First, as opposed to plea cases (where pleas and sentences are generally negotiated between the prosecution and defense and are predominantly a reflection of prosecutorial discretion), the sentencing phase after trials offers a purer reflection of judicial sentencing discretion. Second, in plea cases, defendants are generally required to admit their guilt to the court. Finally, judges are exposed to more case evidence during trials, where witnesses testify, scientific evidence is presented, and cross-examinations are performed.

confidence in guilt -- influence sentence length. Since additional pieces of evidence corroborate other case evidence, more evidence should signal a higher level of confidence in guilt on cases in which convictions were secured at trial. Likewise, judicial perceptions that forensic evidence is more objective and stronger than other evidentiary forms (Saks and Faigman, 2008) may explain why cases with forensic evidence result in longer sentences post-conviction. While these quantitative analyses can conclusively identify sentencing disparities, they cannot fully explain the mechanism by which case evidence is translated into sentence variation. The qualitative analyses provided in Chapters 5 and 6 seek to fill this gap. In Chapter 5, I explored judicial perceptions of various forms of evidence. Judges confirmed that they perceive forensic evidence to be more objective and stronger than witness testimony.⁵¹ Moreover, they expressed an increased confidence in guilt when the cases included scientific evidence. These analyses further confirmed that additional pieces of evidence (which corroborate other case evidence) also increase judicial confidence in guilt. Thus, Chapters 4 and 5 have established, respectively that: (1) cases with forensic evidence and greater quantities of evidence result in longer sentences for convicted defendants, and (2) the existence of inculpatory forensic evidence and greater amounts of evidence available in cases increases judicial confidence in guilt. What remains unknown is whether judges consciously or explicitly use guilt-phase evidence as a justification for sentence determinations in the post-verdict context.

Drawing from these findings in this final analytic chapter, I seek to answer the following question: how, if at all, do judges consider strength of the evidence at

⁵¹ While most judges perceive forensic evidence to be more objective and stronger than non-scientific testimony, this varies among forensic evidence types. For example, judges generally reported that DNA is more reliable than fingerprint comparisons.

sentencing post-trial? In qualitative interviews with judges, I rely on two different strategies to explore this issue. One strategy involves direct inquiry through an interview guide approach and the other involves the use of vignettes. With regard to the former, I asked judges the following questions: Do you consider the strength of the evidence in determining sentence? And do your own perceptions of guilt and/or the accuracy of the jury's verdict factor into your sentencing decisions? These questions were posed at the very end of the interview (and always after the vignettes) to prevent tainting judicial responses to earlier questions. In the first part of this chapter, I report these findings.

While direct inquiry provides a great deal of important insight, this method is only capable of accessing conscious perceptions. In order to access subconscious processes, I presented four written vignettes to judges. These vignettes offer scenarios in which a jury convicts a defendant of either assault or robbery. The vignettes control for many non-evidentiary factors⁵² but vary in the amount and type of evidence available to the court in each scenario. After reading each vignette, judges were asked to sentence the defendant and to describe the rationale by which they arrived at their hypothetical sentences. In the second part of this chapter, I compare and contrast responses to these vignettes. After providing a brief description of each vignette, I report and analyze judicial sentiments relating to evidentiary weight and the hypothetical sentences imposed by judges. Then, I compare and contrast responses between the two aggravated assault vignettes and the two robbery scenarios, analyze differences in sentence length, and conclude by discussing the significance of these findings.

⁵² Some controls include gender and sentencing range. For a detailed discussion of how the vignettes were designed and accompanying rationales, see Chapter 3.

A. Direct Inquiry

I asked judges directly whether their opinions regarding the defendant's guilt, arising from perceptions of case strength, factor into their sentencing decisions. Of the 41 judges in this study, 20 explicitly stated that their perceptions of guilt influence their sentencing determinations. Four additional judges noted that, while they were never faced with a situation where they disagreed with the jury's verdict, doubts regarding guilt would probably influence the punitiveness of their sentences. In contrast, ten judges stated that they do not consider the strength of the evidence or weigh their personal feelings regarding the defendant's guilt in assigning a sentence. A majority of these latter judges maintained that these issues are considerations for the fact-finding jury and thus outside of their purview. Four judges refused to answer the question (one told me emphatically to "move on") and three judges stated that they had never thought about it or were unsure.

Affirmative responses

The 20 judges in this study who reported using guilt perceptions as a factor at sentencing articulated various reasons for considering evidentiary weight. Several emphasized the importance of maintaining their own peace of mind. Having to impose sentence on a defendant, whom they perceive to be innocent, is personally unsettling for these judges; the idea that their actions may place an innocent defendant in prison is difficult to bear. Consequently, when the evidence is weak and confidence in guilt is low, they attempt to ease their consciences by "do[ing] what [they] can" to minimize perceived injustices: "If it's a weak case and I believe that the defendant is innocent, I

sentence to the minimum no matter what his record looks like. I have to sleep at night.”

Another commented: “If I believe the defendant did not do it, I apply rustic justice. I am, after all, human.” And a third said: “The stronger the case, the more calming it is for me to bang them.”

Other judges indicated that they had not consciously considered these potential ramifications. However, they reflectively concluded that their perceptions of evidentiary weight are influential at sentencing, but likely operate under the surface rather than explicitly. For example:

It would probably creep in. I don’t know if I would sit there and consciously say that I am only going to give him two years instead of six because it was pretty weak evidence, but I am just being honest. It would creep in and somehow I would factor it in if I really felt that it was a close call and the guy might be innocent. Wow.

A couple appeared conflicted; while acknowledging that evidentiary considerations may be inappropriate at sentencing, these judges indicated that such considerations are somewhat inevitable:

I am not sure whether it should be, but human nature being the way it is, I can’t get away from that. If I am absolutely convinced that this person committed the crime, I am more content giving a more severe sentence than if I have my doubts. Whether it’s consciously or unconsciously, that does come into play.

In a few instances, judges appeared rattled by this line of questioning and indicated that they had not pondered these issues before. One judge, who had never faced such a conflict during his judicial career, nor consciously considered this scenario, acknowledged the potential impact of its occurrence:

That’s a tough question. I almost think common sense would say “what judge wouldn’t think of that?” How could you not factor that in? You start thinking, “Boy, I probably wouldn’t have convicted if I were on that jury.” What if . . . gosh. I don’t know how it creeps its way into my thought processes to be honest with you.

In addition to considering their personal perceptions regarding evidentiary strength and defendant guilt during the sentencing process, a few judges additionally consider whether or not the jury convicted the defendant of the proper charge. These considerations can cut both ways; perceptions of “over-convictions” may lead to shorter sentences (e.g. the judge believes that the defendant should have been convicted of Assault in the Second Degree but the jury convicted the defendant of Assault in the First Degree, with a more punitive range of sentencing possibilities) while perceptions of under-convictions may result in longer sentences (e.g. the jury acquitted of the top charge, despite a strong evidentiary package and convicted the defendant of a lower level offense): “If I don’t agree with the jury, I consider that in my sentencing. I’m going to sentence the guy like he was convicted of the charges that I felt he should have been convicted of and he is not going to get as much as he would have.” Another stated: “If the defendant is over or under convicted, I account for that at sentencing.”

In addition to their own experiences, a handful of judges discussed the scholarly literature regarding the inherent unreliability of eyewitness identification testimony (Loftus, 1974; Sanders, 1984; Sharps et. al, 2007; Shermer, Rose and Hoffman, 2011). Fearful that a mistaken identification may result in a miscarriage of justice, several claimed, unequivocally, that they impose lower sentences when they are concerned that the identification of the defendant may be inaccurate. For example, one noted: “Defendants always get a lower sentence from me in a one-witness id case.” Another stated: “If I am sure that he did it – there’s DNA, he left his wallet with his name on it at the scene, there are five eyewitnesses, I am going to consider that. The stronger the case, the more time you are going to get.”

Several judges reported that they consider the potential of being reversed on appeal when they determine sentence. For example:

I always consider strength of the evidence because there is going to be an appeal. One of the things I consider is how bulletproof the conviction is going to be. One of the best ways to invite a reversal is to impose the maximum or very serious sentence. If the sentence is less, there's less likelihood of reversal. If it's more, sometimes they come up with all types of excuses to reverse the conviction when they don't like the sentencing. Pure and simple.

Approximately a third of the judges focused on the devastating effects of wrongful convictions in explaining their decision to integrate strength of the evidentiary package into their sentencing decisions. One noted: "Strength of the evidence is very important in determining sentence. It's terrible to be convicted of a crime you did not commit, so if I have photos and medical evidence I give a longer sentence." Another stated: "Strength of the evidence is an especially big deal in sentencing defendants in domestic violence cases. It is unthinkable to be convicted of a crime that you did not commit, so cases with compelling evidence get longer sentences."

A few acknowledged that they felt limited in their ability to correct wrongful convictions. In other words, while lower sentences mitigate damages, it is hardly a complete remedy: "It's not enough, I know, but if I really believe he's innocent, I am going to give the minimum the law permits. It's better than nothing." And another said:

When a guy loses a case and I am comfortable that he is guilty, it is one thing. If the case is so bad that I don't think that the jury should have convicted him it's another. I would be more apt to give probation if I could. The jury verdict goes but I have to speak with my verdict too.

A couple described instances where "their hands were tied" by the jury verdict. Despite concluding that the verdict was legally reached and, consequently, should not be set

aside, the possibility of error weighed heavily on these judges. Therefore, they opt for leniency “just in case.”

When you consider the strength of evidence at sentencing, it’s almost as if you are saying “well, the evidence is not that strong but I have to sentence him even though he might not have done it so I will give him a lesser sentence.” That’s not very intellectually satisfying, although it is what it is. You always have to recognize that you could be making a mistake, which is one reason why the death penalty is such a ridiculous thing.

Finally, sentencing ranges allow for corrective measures with both ceilings and floors. These legally sanctioned options allow judges to inject their own perceptions and beliefs into the process without taking the extreme legal action of setting aside a jury verdict:

Strength of the evidence is a very important factor. Even though you are convicted I still have a range to work with. I can understand that the jury tried to do their job, but if the evidence wasn’t so good I will give him a lesser sentence. If the evidence is totally supportive, then it goes to the other end.

Negative responses

In contrast, approximately a quarter of the judges (n=10) stated that evidentiary weight and confidence in guilt do not play a role at sentencing. Some were emphatic. For example, one said:

I have disagreed with juries and been unhappy. But then I work on myself and say that it is the jury’s call. Weak evidentiary packages do not make a difference to me . . . I have handed out tough sentences after disagreeing with a jury’s guilty verdict.

Another argued: “I don’t care about the evidence. It is irrelevant if the jury found him guilty.” And a third noted: “My personal feelings have nothing to do with sentencing.”

In a few instances, judges were very clear that it would be stepping outside their authority to, in effect, revisit guilt decisions by adjusting sentences accordingly. For instance:

Strength of the evidence doesn't matter. If I am not going to set it aside and the jury has spoken, I've got someone who is convicted and the facts are the facts. That's the way I look at it. I know that different judges disagree and I know that some judges look at strength of the evidence. But my philosophical approach is that once you are found guilty, you are guilty.

A second explicitly referenced the right to a jury trial:

I won't usurp the role of the jury so my sentence is going to be based on the assumption that the defendant is guilty. I am just not going to go back and question the issue of guilt. The jury has told me . . . that's their job. They've said, "whatever the quality of the evidence is, it's sufficient. It's beyond a reasonable doubt, we're satisfied, we're the finders of fact." To me, he's just as guilty as if the evidence is 100% -- if there is such a thing.

Summary

As these data demonstrate, approximately half of the judges acknowledged that they have actually considered or would evaluate the strength of the evidence at sentencing. This is especially true in situations where the judge experiences doubts regarding the defendant's guilt. Judges focused on a variety of reasons for considering evidentiary weight and confidence in guilt at sentencing including maintaining peace of mind, mitigating perceived guilt phase injustices, hedging against potential errors, accounting for over or under convictions, and avoiding reversal on appeal. Only a quarter of the judges (n=10) articulated complete deference to the jury's verdict in situations where they experienced doubt regarding the defendant's guilt. Referencing the defendant's right to a jury trial, these judges cast aside doubts of guilt (and, in some cases, "work on themselves") and avoid considering evidentiary weight at sentencing.

While direct inquiry clearly demonstrates that many judges consider evidentiary weight at sentencing, how these processes unfold during the sentencing phase remains unclear. In fact, some judges specifically stated that evidentiary weight may be

influencing their decisions on a subconscious level. Nonetheless, direct inquiry revealed that approximately half of the judges consciously consider evidentiary weight at sentencing for a host of moral and practical reasons.

B. Vignettes

As the previous section shows, many judges are conscious of the influence of evidentiary weight on their sentencing decisions and candidly shared their views; yet analyses based solely on direct inquiry may limit the findings to conscious thought processes or retrospective reinterpretations. In fact, several judges articulated the possibility that subconscious processes may be influencing their sentencing decisions. Therefore, in this section, judicial responses to the vignettes are analyzed to access subconscious considerations and reflective thought processes (Mah, Taylor, Hoang & Cook, 2014).⁵³

Summaries of the Vignettes

As previously indicated, four different vignettes were utilized in this study. Summaries of these vignettes are provided in this section. The actual vignettes are included in Appendix 3.3.

[INSERT TABLE 6.1 HERE]

⁵³ I employed these vignettes in 22 of the 41 interviews. Vignettes were not included in 19 interviews for a variety of reasons including refusal by the judge (n=11) and time constraints (n=8). Additional explanations for judges declining to participate in the vignette portion of the interview can be found in Chapter 3, Footnote 31.

Scenario A

Scenario A presents a fact pattern describing an aggravated assault in which the evidentiary package contains neither forensic evidence nor eyewitnesses (other than the victim). In Scenario A, the defendant was convicted of stabbing the victim in the leg with a knife (the victim received 12 stitches as a result). The incident occurred in a bar after the victim refused the defendant's order to move to a different stool. The victim identified the defendant to the police, in the bar, ten minutes later. While ten people were present at the time, they all claimed that they did not see the incident and the knife was not recovered (see Appendix, 3.3, Scenario A; see also Table 6.1).

Scenario B

Scenario B presents a fact pattern in which the evidentiary package contains forensic evidence and eyewitnesses (in addition to the victim). In Scenario B, the defendant was convicted of punching the victim in the stomach and stabbing him in the arm with a small knife wherein the victim received eight stitches as a result. The incident occurred at a party after the defendant accused the victim of making advances toward his girlfriend. Three eyewitnesses corroborated the defendant's version of events and the knife was recovered. Forensic analysis confirmed that there was blood on the knife matching the victim's blood type (see Appendix 3.3, Scenario B, see also Table 6.1).

Scenario C

Scenario C presents a robbery fact pattern in which the evidentiary package contains an eyewitness (in addition to the victim) but no physical or forensic evidence. In

this hypothetical, the defendant was convicted of forcibly taking the victim's cell phone and slashing her hand with a box cutter in the process. Following the incident, the defendant fled the scene but was apprehended by the police five minutes later (ten blocks from the incident location). The victim and an independent eyewitness, who observed the incident, identified the defendant. The box cutter was not recovered (see Appendix 3.3, Scenario C, see also Table 6.1).

Scenario D

Scenario D presents a fact pattern in which the evidentiary package contains forensic evidence but no eyewitnesses (other than the victim). In Scenario D, the defendant was convicted of forcibly taking money from the victim at knifepoint, grabbing a chain from her neck, and slashing the victim's arm (the victim received stitches and antibiotics as a result). The incident occurred on the street as the victim was walking home from work. The victim identified the defendant to the police within twenty minutes and eight blocks from the incident location. The knife was recovered at the scene. Subsequent testing revealed that it contained fingerprints matching the defendant's prints (see Appendix 3.3, Scenario D, see also Table 6.1).

Non-Evidentiary Features

As Chapter 3 describes in detail, the vignettes were designed to control for a variety of non-evidentiary factors (i.e. factors that do not relate to the weight of the evidence). By reducing the non-evidentiary differences between the scenarios, evidentiary effects may be better observed. To that end, in each hypothetical scenario all

defendants were first time offenders, and were convicted by a jury of a class B violent felony⁵⁴ with a sentencing range of 5 to 25 years. Defendants were all described as male of unspecified age, who did not confess. Further, the victims in all of these scenarios received minor injuries, via a knife or box cutter, for which they received stitches. Race was not provided in the written descriptions. Yet, due to practical challenges in designing these hypotheticals⁵⁵ and the unique features described in different scenarios, not all non-evidentiary case characteristics could be controlled. For this reason, I asked judges to “walk me through the process” by which they arrived at their sentences; this unguided discussion provided rich detail on the factors that motivated sentence severity as well as insight into the varying rationales and analyses that drove these motivations. In this section, I briefly describe the non-evidentiary differences between the vignettes. This description is important in that it highlights perceived distinctions in crime severity that help to explain some of the observed differences in sentence lengths among the four scenarios.

Scenarios A and B are aggravated assaults while Scenarios C and D are robberies. Despite the fact that all four vignettes are Class B felonies in New York with the same sentencing ranges, judges perceived the crime of robbery to be more serious than the crime of aggravated assault since robbery involves the taking of property. Further, the two robbery scenarios were premeditated street crimes. For these reasons, judges perceived offense seriousness to be more severe in Scenarios C and D than in Scenarios

⁵⁴ In New York, felonies are categorized by Classes A through E; A is the most serious class and E is the least serious. A variety of crimes are classified as B violent felonies. Among others, these include certain assaults, rapes, robberies and kidnappings. B violent felonies in New York State have a sentencing range of 5 to 25 years (see N.Y. Penal Law, Section 70).

⁵⁵ See Chapter 3 for a detailed description of the construction of the vignettes.

A and B. Given these significant differences, comparisons made among the vignettes was conducted within crime type.

A few differences between scenarios within the same crime type should be noted from the outset. First, judges perceived the defendant in Scenario A to be completely unprovoked. In contrast, nearly half of the judges in Scenario B focused on the victim's behavior in evaluating the blameworthiness of the defendant; most of these judges concluded that the victim instigated the defendant and that the crime was one of passion. Thus, judges found the defendant in Scenario B to be less blameworthy than the defendant in Scenario A.

Similarly, judges perceived differences in crime severity between Scenarios C and D. By far, Scenario D was viewed by judges as the most serious and violent (the defendant grabbed the victim's chain off of her neck and the victim was slashed, even after complying with the defendant's demands). While Scenario C was also viewed as violent, the victim in that scenario was injured while attempting to grab her phone back – a fact perceived by several judges to be mitigating.⁵⁶

In the sections that follow, I begin by reporting judicial considerations and sentences for each vignette and conclude this section by comparing and contrasting reactions to each of these scenarios, within crime type.

Analyses of four scenarios

Judges were handed a written copy of each vignette, one at a time, and asked to read the vignette and render a sentence. In describing the analyses that led to their

⁵⁶ The four vignettes are available in Appendix 3.3.

sentences, judges were not prompted to discuss any specific sentencing considerations.

To the contrary, without direction, I asked them to describe all of the factors and considerations that contributed to their sentences.

Scenario A

Evidentiary Weight

Without any prompting, more than half of the judges (12 of 21) discussed evidentiary weight, among other factors, in explaining the sentences they imposed for the defendant in Scenario A. In discussing the evidence, virtually all of these judges focused on the lack of eyewitnesses (other than the victim) and viewed it as an evidentiary weakness. They interpreted its importance and significance in different ways. For a few judges, the absence of this evidence presented reliability concerns; without eyewitnesses, the victim's testimony was left uncorroborated. This relegated Scenario A to the weak status of a one-witness identification case, where concerns of witness reliability are paramount (Connolly & Gordon, 2011; Shermer, Rose & Hoffman, 2011). It is noteworthy that judges sometimes explicitly linked their reliability concerns to a willingness to be lenient in sentencing. As one judge stated, "It's a one person id. I would go as low as I could."

Others were focused on the victim's credibility and considered whether the victim's version of events was consistent with an absence of eyewitnesses:

I have issues with no one else having seen it even though I appreciate that it's a bar and no one was paying attention. I tend to think that there could have been loud words at some point. Somebody would have at least looked in that direction, even if they did not see the stabbing. They should have seen an altercation, or heard an altercation, so I have some doubts.

A second stated, “You would think that there would be an independent witness, someone in the bar. There is no corroboration for the victim’s story.” Notably, two judges concluded that the lack of eyewitnesses was reasonable in this situation and, consequently, did not undercut the victim’s credibility. For example: “He stabbed him in the leg. Who is looking underneath the bar”? These quotes demonstrate that certain judges employed their common sense and life experiences to assess the victim’s credibility and the plausibility of the lack of eyewitnesses given the circumstances of the case. These credibility assessments are jury functions. These judges clearly assumed this function by considering the lack of eyewitnesses in light of their own common sense (e.g. who is looking underneath the bar?) and life experiences.

Judicial consideration of these evidentiary factors is problematic because the jury in Scenario A already assessed the victim’s credibility and reliability, and found the defendant guilty. Whether or not eyewitnesses corroborated the victim’s version of events does not speak to the blameworthiness of the defendant or the severity of the offense. Instead, it relates to perceptions of case strength; these considerations pertain to whether the victim’s uncorroborated testimony is sufficient to establish the defendant’s guilt. Indeed, eyewitnesses provide probative testimony and supportive identification evidence at trial (Loftus, 1974; Skolnick & Shaw, 2001). Eyewitnesses can corroborate a victim’s version of events and bolster the victim’s credibility. Conversely, the lack of eyewitnesses may be perceived as a weakness in the case. These are precisely the calculations that the Supreme Court ruled should be left to the purview of the jury. They relate to guilt determinations and are not valid sentencing factors (*Alleyne v. U.S.*, 2013; *Apprendi v. New Jersey*, 2000; *U.S. v. Gaudin*, 1995).

The assumption of a fact-finding role is likewise evident from judges' focus on the lack of physical or forensic evidence in the case. One third (n=7) considered the fact that the knife was not recovered in determining sentence. Again, the missing knife was analyzed in opposing ways. While each of the seven found the missing knife to represent a weakness in the evidentiary package, two judges reasoned that there was a plausible explanation for its absence. In particular, one argued: "The fact that the knife was not recovered is the only weakness in the people's case. [But] there are many reasons why the knife might not have been recovered since there is a lapse in time from when the call was made to when the police arrive." In contrast, another reached a different conclusion: "They never found the knife . . . that seems a little bit odd . . . with so little evidence, no more than the minimum."

Nearly half of the judges (9 of 21) either did not mention the evidence at all (n=5) or mentioned the weakness of the evidence but then stated that they did not consider evidentiary weight when reasoning through their respective sentence calculations (n=4). However, two of the judges who affirmatively mentioned that case strength did not matter, since the jury already rendered a verdict, went on to discuss evidentiary weaknesses in this scenario. For example, one stated: "It doesn't matter to me that the ten individuals claim that they didn't see the defendant, or the knife was not recovered . . . I'd have to defer to the jury on that." A different judge analyzed the defendant's conduct in assessing whether his actions were indicative of guilt, but quickly noted that these factors should not be considered during sentencing: "The defendant didn't flee the scene and denied guilt on all of this. But this is ultimately irrelevant because he's been convicted." Thus, despite assertions that the strength of the evidence is not a sentencing

concern, these two judges nevertheless highlighted perceived evidentiary weaknesses in the case and definitively referenced facts relating to proof of guilt in determining a sentence.

Sentences Imposed

Scenario A represents the weakest evidentiary package of all of the vignettes; it is the only vignette in which the evidence does not include eyewitnesses or forensic evidence. As Table 6.2 indicates, sentences ranged from probation to up to 10 years in prison, with an average sentence length of 3.7 years. An analysis of the hypothetical sentences reveal differences in sentence length between judges who discussed the strength of the evidence in describing their sentencing processes (n=12) and those who did not consider evidentiary weight (n=9). Of the 12 judges who discussed the lack of eyewitnesses and/or the failure of the police to recover the knife, the average sentence length for Scenario A was 2.8 years, whereas those who did not consider evidentiary weight sentenced the same defendant in the same circumstances to nearly twice as long (at approximately 4.95 years of incarceration, on average). Further, the three most punitive sentences rendered in Scenario A (i.e. up to 10 years, 9 years and 7 years) were all imposed by judges who did not consider evidentiary weight.

[INSERT TABLE 6.2 HERE]

The combination of judicial rationales and hypothetical sentences imposed in Scenario A suggest that approximately half of the judges considered evidentiary weight at sentencing. In fact, those who considered the weak evidentiary package explicitly

rendered hypothetical sentences that were considerably shorter in duration than did judges who avoided considering evidentiary weight. These findings suggest that some judges impose lower sentences when they perceive the evidentiary package as weak, despite a legally rendered jury conviction.

Scenario B

Evidentiary Weight

Two thirds of the judges (n=14) considered evidentiary weight when describing how they calculated their sentences. This proportion is higher in Scenario B than in any other scenario. Scenario B represents the strongest evidentiary package of all of the vignettes; it is the only vignette that includes forensic evidence and eyewitnesses. In general, judges were impressed by the evidence in this case and felt that it was worthy of mention. In fact, virtually all of the judges who discussed evidentiary weight commented on the strength of the evidence; the combination of eyewitnesses and forensic evidence instilled strong confidence in guilt. For example: “This is an open and shut case. There is not much you can argue with here. Although they didn’t know each other, the eyewitnesses all corroborated the complainant’s version and the knife was recovered.”

Several judges explicitly indicated that their sentence were more severe due to the strength of the evidence: “This is a strong case. The matching blood corroborates the fact that the knife was used, so I gave him 8 years.” Another noted: “You have independent witnesses and you have a weapon recovered – clearly that factors in.”

A handful of judges juxtaposed the strong evidentiary features in this vignette against their perceptions that this case was not that serious, due in part to the victim’s

blameworthy behavior⁵⁷ “While the victim may have had it coming, three eyewitnesses corroborated the story and the case has forensic and physical evidence – I can’t ignore that.” Another said: “It’s all over some girl but it’s a really good case for the prosecutor. Between the knife and the blood and the witnesses, at least we know he really did it.” While these judges minimized the severity of the offense based on victim precipitation and motivation, these perceptions were juxtaposed against perceptions of evidentiary strength. Confidence in guilt seemed to justify imposition of a more punitive punishment, despite perceptions that the offense severity level was low.

In total, one third of the judges (n=7) did not refer to the evidence when they described their sentencing rationales. Unlike all other scenarios, none of these judges discussed the irrelevance of evidentiary weight – they simply did not mention the evidence at all.

SENTENCES IMPOSED

The sentences imposed in Scenario B ranged from alternatives to incarceration to 9 years in prison. On average, judges sentenced the defendant in Scenario B to 4.5 years in prison (see Table 6.2). Distinct differences in sentence length are evident between judges who discussed evidentiary weight (n=14) and those who did not (n=7); judges who considered the strong evidentiary package in Scenario B imposed sentences, on average, a third higher (mean sentence length = 5.1 years) than judges who did not consider evidentiary strength (mean sentence length = 3.4 years). It is noteworthy that while a total of 7 of the 21 judges sentenced the defendant to a period of incarceration

⁵⁷ For a more complete discussion of judicial impressions of crime severity for this scenario, see pages 156-159.

above the mandatory minimum, all of the judges who did not consider evidentiary weight imposed sentences below or at the minimum (with the exception of one judge).

Viewing the hypothetical sentences in conjunction with judicial comments reveal that judges were not only impressed by the strong evidentiary package, but, also, that these perceptions resulted in more punitive sentences. More than a third of the judges explicitly and unequivocally stated that they would impose harsher sentences for the defendant in Scenario B because of the strength of the evidence. A comparison in average sentence length between those judges who did and did not discuss the evidence shows significant differences; sentences were significantly longer when evidentiary weight was considered. These results provide strong support for the conclusion that judges impose longer sentences when the evidentiary package is strong.

Scenario C

EVIDENTIARY WEIGHT

More than half of the judges (13 of 21) spontaneously discussed evidentiary weight while describing their rationales in determining sentence and eight judges either did not mention evidentiary factors (n=6) or specifically noted that the weight of the evidence is not a sentencing consideration post jury verdict (n=2).

Judges who considered evidentiary weight did so for different reasons. Judges mentioned both strengths (i.e. a corroborating eyewitness) and weaknesses (i.e. box cutter and cell phone not recovered) in the evidentiary package. More than half who considered evidentiary weight (n=7) noted that the lack of physical evidence weakened the case; but, again, the importance attributed to this fact was analyzed in diverse ways. For example,

one stated: “Not only wasn’t the box cutter recovered, but the cell phone was not recovered either. That’s a little unsettling.” In contrast, a different judge evaluated the missing box cutter and concluded that the police’s failure to recover it was reasonable under the circumstances and consistent with other case facts: “Box cutters are cheap. [The defendant] might have decided to toss it and the police are not going to engage in an extensive search, especially since it’s ten blocks away and it’s just a slash. How much time are they really going to spend looking for it?” Despite varying judicial speculations regarding the reason for the police’s failure to recover these items, all seven concluded that the lack of physical evidence weakened the case and three explicitly stated that they would impose a lower sentence as a result: “You would think at least some evidence would have been recovered. Five year minimum sounds about right for this garbage.” Moreover, the absence of physical evidence precluded forensic examination in a laboratory, which further weakened the evidentiary package for a few judges: “There is nothing recovered. Nothing to test. Weak case.” This is consistent with prior studies that conclude that the courts place a great deal of confidence in forensic science (Saks & Faigman, 2008, p. 150; see also Edmond & Roque, 2012).

In contrast, more than a third ($n=8$) of the judges discussed the presence of an independent eyewitness in the case and explained that this fact strengthened the evidentiary package: “You got the defendant being identified in a lineup by someone who was right across the street – that’s good.” Another argued: “An independent witness came forward with no reason to lie. It’s a strong case. I am sending him to jail.” In contrast to one-witness identification cases (e.g. Scenario A), concerns regarding the inherent unreliability of eyewitness testimony and the possibility of a mistaken

identification (see Connolly & Gordon, 2011; Shermer, Rose & Hoffman, 2011) were soothed by the presence of a corroborating eyewitness.

Less than half of the judges (8 of 21) either did not refer to the evidence (n=6) in describing their sentencing considerations or affirmatively stated that the weight of the evidence is irrelevant at sentencing (n=2). These latter two judges, however, relied on evidentiary weight in their rationales, despite formerly claiming that guilt phase evidence is not a valid sentencing criterion. In fact, they did not mention *any* non-evidentiary considerations in their discussion of Scenario C. For example, one judge, who imposed a sentence significantly below the average for judges who did not consider evidentiary weight (i.e. 5 years compared to the 7.6 year average), noted: “The conviction is the conviction. I am able to separate the absurdity that someone who just used a box cutter for dismantling boxes at his job would then use it to rob a cell phone – ridiculous” Another judge, comparing this vignette to Scenario A, stated: “At least here you have a lineup with an eyewitness as opposed to ten witnesses who saw nothing. But the jury made a decision and convicted, so none of that affects me.” Notably, this judge imposed a sentence of “up to 10 years,” which is considerably higher than the average sentence imposed in Scenario C. Despite emphatically asserting that evidence does not matter, these judges focused on case-specific evidentiary features and did not mention legally relevant sentencing factors, such as case seriousness or criminal history, in their discussions.

SENTENCES IMPOSED

Sentences imposed for Scenario C ranged from a split sentence (i.e. six months in jail and five years of probation) to 15 years in prison, with an average of 6.5 years. There

were significant differences in sentence length between judges who considered evidentiary weight (average = 5.8 years) and those who did not (average = 7.6 years); those who discussed the evidentiary package sentenced defendants to significantly shorter sentences. Of the 13 judges who considered evidentiary weight, seven discussed the missing knife and/or cell phone⁵⁸ and eight judges noted the presence of a corroborating eyewitness.⁵⁹ Sentences varied among these 13 judges. The average sentence length for those who discussed missing physical evidence was 5.1 years, compared to an average sentence length of 6.6 years for those who only focused on the presence of a corroborating eyewitness.

These numerical sentences, viewed in conjunction with judicial responses, are suggestive of several factors. First, the lack of physical evidence appears to be more influential on sentence length than the presence of the eyewitness. In fact, the average sentence length for judges who discussed the eyewitness but not the missing evidence was virtually identical (6.6 years as opposed to 6.5 years) to the overall average sentence imposed by all judges in Scenario C. This suggests (similar to patterns observed in the quantitative analyses), that the presence of eyewitnesses, while reassuring for some judges (as evidenced by their remarks), is not associated with longer sentences post-conviction. In contrast, judges focused on the lack of physical evidence imposed the shortest sentences in Scenario C; thus, these findings suggest that physical evidence (or the lack of physical evidence) is more influential at sentencing than witness-based evidence for some judges.

⁵⁸ Five mentioned only the missing physical evidence and two mentioned both the lack of physical evidence and the presence of a corroborating eyewitness.

⁵⁹ Six focused exclusively on the eyewitness and two noted both the presence of an eyewitness and the absence of physical evidence.

*Scenario D***EVIDENTIARY WEIGHT**

Approximately half of the judges (11 of 21) openly discussed evidentiary factors during the course of describing their sentencing rationales. These judges focused on the strength of the evidentiary package. While there were no eyewitnesses other than the victim, most judges were not concerned about the identification of the defendant (three explicitly stated that eyewitnesses were not necessary here and six noted the lack of eyewitnesses but did not elaborate) because of the presence of forensic evidence. Specifically, the defendant's fingerprints recovered on the knife instilled a high degree of confidence that the victim's identification of the defendant was correct. In fact, all of the judges who discussed evidentiary weight discussed the presence of fingerprint evidence:

You have fingerprints on the weapon, which were recovered at the scene, so that eliminates the issue of identification. I am always concerned about that, whether the witness is correct. When you see DNA or fingerprint corroboration for the witness, you feel comfortable. You don't want to sentence somebody who is innocent.

This comment, echoed in similar fashion by several judges, exemplifies the heavy weight afforded to fingerprint evidence and is consistent with prior studies that establish that the courts place a great deal of confidence in forensic science (Saks & Faigman, 2008, p. 150; see also Edmond & Roque, 2012). Yet, this comment, that the recovered fingerprints "eliminate the issue of identification," is concerning; this is because research shows that forensic evidence is not as conclusive and objective as generally perceived. In fact, scholars have expressed that forensic science practices often involve speculative opinions and questionable scientific techniques (Cole, 2009; Dror et al., 2011; Dror, 2013;

Edmond & Roque, 2012; Saks & Faigman, 2008). For example, Dror (2013) addressed fingerprint processes specifically and found that subjectivity and bias are inherent in certain latent fingerprint analyses. While the fingerprints recovered in this case clearly corroborate the victim's identification of the defendant and logically increase perceptions of case strength, several judges unequivocally concluded that the identification issue was resolved (e.g. "that eliminates the issue of identification") by the mere presence of this forensic evidence.

Other circumstances surrounding the identification were reassuring to several judges. In fact, a few judges noted that the identification was made only 20 minutes after the crime occurred and 8 blocks from the crime scene:

The police arrived within minutes. The identification was in close proximity not only to where the crime was committed, but also in respect to when the crime actually occurred. You weren't asking someone to wait two months and say, "Hey, can you point out this stranger again."

Notably, the identification of the defendant in Scenario A was conducted at the scene of the crime (i.e. the bar), as soon as the police arrived. In fact, the defendant in Scenario A never left the bar (according to the victim). Not one judge characterized the victim's identification of the defendant in Scenario A as strong. To the contrary, judges focused on the lack of eyewitnesses and physical evidence. Yet, the facts surrounding the identification in Scenario A are stronger than those in Scenario D, since the defendant remained at the crime location (i.e. the bar) which provided the victim with extended opportunity to observe the defendant. Moreover, the identification took place at the scene of the crime. While the conclusion (i.e. in Scenario D), that an identification conducted twenty minutes after and eight blocks from the crime did not raise reliability concerns

may be reasonable, judicial silence on this issue (by these same judges) in Scenario A is puzzling. This raises the possibility that the reassuring qualities of the forensic evidence present in Scenario D and the lack of evidence in Scenario A colored perceptions of other case evidence. Additionally, and perhaps more significantly, it is another example of judges assuming a fact-finding role and usurping the jury's verdict.

While eleven judges mentioned evidentiary weight, time spent discussing evidentiary characteristics was less in this scenario than any other. Instead, the focus was on the heinous and serious nature of the offense. Reviews on evidentiary strength were positive (and focused on the fingerprints) but they were not nearly as overwhelming as those in Scenario B (which contained both eyewitnesses and forensic evidence).

Ten judges either did not mention case evidence at all (n=7) or affirmatively stated (n=3), without prompting, that guilt phase decisions are within the purview of the jury and irrelevant at sentencing. As one noted: "I really don't care about the questions about the evidence. Whether the evidence is strong or not, the jury found the defendant guilty. It's on them – it's not my responsibility." The vast majority of these ten judges focused predominantly on non-evidentiary factors that they perceived as aggravating or mitigating.

SENTENCES IMPOSED

Sentences imposed for Scenario D were more severe than sentences imposed in all of the other vignettes. In this case, they ranged from 5 to 15 years in prison, with an average of 8.4 years. With the exception of three judges, all other judges imposed their most punitive sentences for Scenario D. Two thirds of the judges (n=14) imposed

sentences in excess of the mandatory minimum – more than in any other scenario. With the exception of one judge, who indicated that he would probably impose a five-year sentence, but may consider a little less if the probation report was favorable, all others indicated that they would not sentence the defendant to a period shorter than five years.

A review of sentence lengths, between those who considered the evidence and those who did not reveal virtually identical means (i.e. 8.4 years compared to 8.5 years). Considering the evidentiary package, the absence of a distinction between those who did and did not consider the evidence makes sense. While forensic evidence instilled confidence in the identification, other case facts were based exclusively on the testimony of the victim. The vast majority of judges who discussed the evidence noted the strength of the fingerprint evidence; yet comments regarding overall case strength were not nearly as positive as in Scenario B, where judges who considered the evidence imposed significantly longer sentences. This scenario also differs from Scenarios A and C in that neither of those scenarios contain forensic evidence; Scenario A relies solely on the victim and Scenario C relies on the victim and an additional eyewitness. While the eyewitness in Scenario C provided corroborative evidence, a third of the judges (who rendered, on average, the lowest sentences in that scenario) were disturbed by the missing knife. In contrast, the vast majority of judges in Scenario D were not affected by the lack of eyewitnesses; a few even explicitly expressed that the presence of forensic evidence rendered eyewitnesses superfluous. These factors provide potential explanations for why, unlike Scenarios A and C, the mean of judges who considered the evidence was virtually identical to the mean of judges who did not.

In sum, despite the lack of eyewitnesses, judges characterized the evidentiary package in Scenario D as strong. In fact, the fingerprint evidence was considered reliable identification evidence for the vast majority and indisputable proof for several. Moreover, the presence of forensic evidence negated concerns relating to the lack of eyewitnesses. These findings provide additional support (in addition to those established in Scenario C, Chapter 4, and Chapter 5) for the conclusion that forensic evidence is more influential at sentencing than eyewitness testimony.

Comparisons Between Scenarios A & B

Comparisons observed between Scenarios A and B strongly suggest that perceptions of evidentiary weight heavily influenced sentencing decisions in these scenarios. This influence is evident by comparing verbalized rationales for the vignettes and the hypothetical sentences imposed. In this section, I will first compare judges' comments regarding evidentiary weight, followed by a comparison of sentence severity.

Two thirds of the judges specifically mentioned that the evidence in Scenario B was very strong. Scenario B includes compelling forensic evidence; a knife containing blood that matched the victim's blood type was recovered. Additionally, three eyewitnesses corroborated the victim's version of events. Not surprisingly, perhaps, as Scenario B included both types of evidence, judges were more likely to comment on the evidence in Scenario B than in any other scenario. Judges stated that Scenario B was an "open and shut case" and a "good case for the prosecutor." Several emphasized a high confidence level in the defendant's guilt with regard to Scenario B (e.g. one judge stated: "At least we know he did it"). In contrast, judges repeatedly mentioned the evidentiary

weaknesses present in Scenario A. Some were disturbed by the lack of independent eyewitnesses, while others were concerned that the knife was not recovered. Several other judges discussed the overall lack of evidence and emphasized that the victim's testimony was uncorroborated. The stark differences in perceptions of evidentiary strength between these scenarios is clear; the vast majority of judges were impressed and reassured by the evidence in Scenario B and underwhelmed and unsettled by the lack of evidence in Scenario A.

Numerical sentences imposed by judges suggest that evidentiary weight influenced sentence severity. Indeed, sentences imposed by judges for Scenario A were lower, on average, than sentences imposed by judges for Scenario B (mean sentence length for Scenario A was 3.7 years compared to the mean sentence for Scenario B, which was 4.5 years). This difference is especially noteworthy because close to half of the judges perceived the defendant to be *less* blameworthy in Scenario B, due to contributory negative behavior on the part of the victim; thus, we might expect the average sentence to be higher in Scenario A. More significant, however, are the differences between judges who considered evidentiary weight compared to those who did not. The average sentence imposed by judges who discussed the lack of evidence in Scenario A was 2.8 years compared to an almost double mean sentence length of 5.1 years for those who considered the evidence in Scenario B.⁶⁰ These numbers reveal opposing patterns between the vignettes. In Scenario A, those who discussed the weak evidentiary package sentenced the defendant to shorter sentences than those who did not, suggesting that the lack of evidence decreased sentence severity for these judges. In

⁶⁰ 11 of 12 judges who considered the evidence in Scenario A also considered it in Scenario B.

contrast, those who considered the strong evidence in Scenario B sentenced the defendant, on average, to longer sentences than those who did not consider the evidence, suggesting that these judges were influenced by the strength of the evidentiary package. Moreover, only three judges in Scenario A indicated that they would sentence the defendant above the mandatory minimum, compared to seven judges in Scenario B. Further, while 13 judges indicated that they would sentence below the mandatory minimum for Scenario A, if permitted to do so, this was true of only 9 judges regarding Scenario B (see Table 6.2). Thus, despite the unprovoked nature of the attack in Scenario A and perceptions of victim culpability by nearly half of the judges in Scenario B, judges imposed higher sentences in Scenario B. These findings are particularly illustrative since Scenario A is the only scenario (of the four vignettes) that contains neither eyewitnesses (other than the victim) nor forensic evidence, while Scenario B is the only vignette with both independent eyewitnesses and forensic evidence.

Comparisons Between Scenarios C & D

Comparisons between Scenarios C and D suggest that variations in judicial perceptions regarding evidentiary type influence the severity of sentences imposed. Specifically, the presence (or absence) of forensic evidence impacts sentence severity while eyewitness testimony is not influential.

In Scenario C, judges commented on the missing physical evidence and on the presence of an eyewitness. While more than a third of the judges verbally indicated that they were reassured by the presence of the eyewitness (who corroborated the victim's testimony), this reassurance is not reflected in their sentences. Specifically, the mean

sentence length of those who positively discussed the eyewitness but not the missing physical evidence was virtually identical (6.6 years compared to the mean of 6.5) to the overall mean sentence length for all judges in Scenario C. In contrast, those judges who discussed the missing knife had a mean sentence length of 5.1 years – considerably lower than the mean. The lack of physical or forensic evidence decreased sentence severity for these judges, while the presence of the eyewitness appears not to be influential.

In Scenario D, approximately half of the judges discussed the evidence; the vast majority noted the absence of eyewitnesses and all referenced the presence of fingerprint evidence in the case. In this scenario, judicial comments reveal the strong influence of forensic evidence (e.g. one judge stated that fingerprints “eliminate the issue of identification”). In contrast, several judges indicated that eyewitnesses were not necessary given the presence of forensic evidence.

While judges in Scenario C, who considered the missing knife, imposed significantly lower sentences than the mean for that scenario, the mean sentence length for all of the judges in Scenario D was virtually identical. With the exception of Scenario B, which contained both forms of evidence, Scenario D is the only scenario in which the absence of an evidentiary form (i.e. eyewitnesses) did not result in lower sentences for judges who considered evidentiary weight in their sentencing rationales. Arguably, the strong probative value and perceived objectivity of fingerprint evidence compensated for the lack of eyewitnesses. These results suggest two concepts: (1) even though judges speak positively of corroborating eyewitnesses, these sentiments do not influence sentence length post-conviction and (2) forensic evidence increases confidence in guilt

and compensates for other perceived evidentiary weaknesses, even after the guilt phase is over.

*Within Judge Comparisons*⁶¹

Individual judges varied in their consistency regarding whether or not they considered evidentiary weight in their analyses of the four vignettes. Six of the judges consistently analyzed the evidence in all four vignettes. Similar to the general patterns, all six of these judges sentenced the defendant in Scenario B (strongest evidentiary package but least serious offense) to a longer sentence than the defendant in Scenario A. Further, during the direct inquiry portion of the interview, all six of these judges indicated that evidentiary strength influences their sentencing determinations. In contrast, only three judges avoided considering the evidence altogether during the vignette portion of the interview. Consistent with the general pattern among judges who did not consider evidentiary strength, two of the three imposed longer hypothetical sentences for the defendant in Scenario A than for the defendant in Scenario B. Further, all three of these judges indicated that they do not consider evidentiary strength in their sentencing decisions during the direct inquiry portion of the interview.

In contrast, the remaining judges considered the evidence in certain vignettes but did not consider it in others. Further, three judges, who stated that they never consider evidentiary strength at sentencing, made reference to the evidence in some of the vignettes. This finding suggests that the influence of case strength on sentencing

⁶¹ Of the 20 judges who did not participate in the vignette portion of the interview, nine judges indicated, during direct inquiry, that they consider evidentiary weight at sentencing, four affirmatively stated that they do not consider evidentiary weight, four refused to answer the question, and three did not know or were unsure.

decisions may operate subconsciously for certain judges. Further, of all of the judges who considered the evidence in both Scenarios A and B ($n = 11$), nine imposed longer hypothetical sentences in Scenario B than Scenario A. In contrast, of the six judges who did not consider the evidence in either Scenario A or B, four imposed longer sentences in Scenario A. These findings are consistent with general patterns. Overall, the results suggest that a majority of judges are inconsistent in their consideration of evidentiary weight at sentencing. Further, some judges may be unaware that evidentiary strength plays a role in their sentencing decisions.

Conclusions

The analyses described above reveal that a small majority of judges consider evidentiary weight during sentencing, post jury-verdict. These analyses are disturbing because sentencing judges who re-evaluate whether the elements of the crime have been proven beyond a reasonable doubt are essentially usurping the role of the jury in violation of constitutional mandates (*Alleyne v. U.S.*, 2013; *Apprendi v. New Jersey*, 2000; *U.S. v. Gaudin*, 1995).

In the first section of this chapter, I discussed judicial responses to direct questions regarding whether judges consider evidentiary weight at sentencing and whether confidence in guilt plays a role in sentencing decisions post-conviction. Approximately half of the judges affirmatively acknowledged that they consider the strength of the evidence at sentencing. Further, confidence level in guilt influences sentencing decisions; judges tend to impose lower sentences when they are unsure about the defendant's guilt.

Likewise, the vignette results reveal that most judges consider evidentiary weight at sentencing and that these considerations actually influence the sentences they impose. First, cases with greater amounts of evidence (e.g. Scenario B) result in longer sentences than cases with less evidence (Scenario A). This conclusion is evident not only in the numerical sentences imposed by judges, but, also, in the comments they made in explaining their sentences. Judges felt confident in the correctness of the verdict in Scenario B, given the strong and voluminous evidentiary package. Conversely, they were unsure about the verdict in Scenario A, since the case was based on the uncorroborated testimony of the victim and there was no physical or forensic evidence in the case. Further, similar to prior studies (Peterson et al., 1987, 2010, 2013), the findings show that the type of evidence in a case influences sentence length. Specifically, forensic evidence instills confidence in guilt for most judges. Conversely, the absence of physical or forensic evidence has the adverse effect. For example, Scenarios A and C did not have forensic or physical evidence. In these two scenarios, judges who considered the evidence sentenced the defendant to shorter sentences than those who did not consider the evidence. In contrast, while Scenario D contained fingerprint evidence, there were no eyewitnesses present to corroborate the victim's version of events. Yet, there were no differences detected in the average sentence imposed between those judges who did and who did not consider the strength of the evidence.

These findings confirm and explain patterns observed in the quantitative analyses described in Chapter 4. Forensic evidence increases confidence in guilt; thus, judges feel more secure in imposing longer sentences when cases contain forensic evidence.

Likewise, cases with more evidence are reassuring to judges; absent concerns over

sentencing an innocent defendant to an undeserved penalty, judges feel confident to impose longer sentences. Thus, evidentiary weight, operating through judicial confidence in guilt, influences sentencing decisions post-conviction.

CHAPTER 7. CONCLUSION

Sentencing disparities are a widespread problem that plague the American criminal justice system. All too often, similarly situated defendants, who are convicted of similar crimes and who possess similar criminal histories, are sentenced to widely disparate punishments. History has shown that the discretionary powers accorded to judges are a major source of these disparities (Frankel, 1972; Lynch, 2007-2008; Lynch, 2009). Thus, understanding the varied factors that influence judicial sentencing decisions is vital to illuminate and limit these harmful consequences.

For the last few decades, scholars have directed considerable attention toward studying the extralegal influences that lead to sentencing disparities. Much of this literature convincingly establishes that defendant and victim sociodemographic characteristics, such as race, gender, age and socioeconomic status, influence sentencing decisions in ways that benefit whites and disadvantage people of color (Albonetti, 1997; Baumer et al., 2000; Baldus, Pulaski & Woodworth, 1983; Mustard, 2001; Osbourne & Rappaport, 1985; Rodriguez et al., 2006; Steffensmeir & Demuth, 2000). More recently, scholars have discovered how characteristics of local courts, within which defendants are sentenced, likewise affect sentence severity and thus perpetuate disparate outcomes (Johnson, 2005; Kramer & Ulmer, 2006). Yet despite this substantial body of scholarly literature, to date, very little attention has been directed toward the influence of evidentiary case characteristics on judicial sentencing decisions in convictions secured through criminal trials. Further, there has been almost no qualitative focus exploring judicial perceptions regarding the role of evidentiary weight at sentencing after a jury verdict.

The lack of scholarship in this area leaves a notable void in our understanding of the extralegal factors that influence judicial sentencing decisions. Indeed, the question of whether or not judges consider strength of the evidence at sentencing post conviction has widespread implications. Evidentiary weight, whether strong or weak, varies among cases and the types and quantities of evidence available in a case are important determinants of whether the defendant will be charged and convicted. In plea cases, evidentiary weight is influential in determining the charge and sentence, since probability of conviction impacts plea negotiations (Bushway, Redlich & Norris, 2014). However, in the post trial arena, judicial consideration of guilt phase evidentiary weight is extralegal and may hamper the imposition of fair and equitable sentences. At this phase, the defendant has already exercised his or her right to a trial and been found guilty. The type and strength of evidence used by the fact finder to establish guilt is no longer relevant to case processing decisions. Post trial sentencing considerations should focus on legally relevant factors such as seriousness of the offense, defendant's criminal history, and other aggravating and mitigating factors.

The type and amount of inculpatory evidence gathered by law enforcement to establish guilt is unrelated to punishment objectives. Nor is it indicative of defendant blameworthiness, the potential for recidivism, or other valid sentencing concerns. Instead, it is a product of the amount of evidence available for collection in the case combined with law enforcement efforts and success in collecting and securing that evidence. The idea that punishment may be influenced, after conviction, by whether the state was able to garner specific types or quantities of evidence against a particular defendant is

problematic. This is because a defendant legally convicted based on a weaker evidentiary package is no less guilty than a defendant convicted based on strong evidence.

In accordance with constitutional mandates, a defendant is entitled to have a jury determine whether all elements of the crime have been proven beyond a reasonable doubt (*U.S. v. Guadin*, 1995; see also *Apprendi v. New Jersey*, 2000). The right to a trial by jury is an essential part of the Bill of Rights. Judges who reexamine the evidentiary package to render a sentence usurp the jury's authority and assume a fact finding role in violation of constitutional principles and Supreme Court judgments (*U.S. v. Guadin*, 1995; see also *Apprendi v. New Jersey*, 2000). While it is natural for judges, post jury verdict, to feel more confident in the conviction when they perceive the evidentiary package to be strong, judges are overstepping their authority when they undermine the finality of jury decisions. Defendants will benefit from this extralegal influence when the evidentiary package is weak, while others suffer longer sentences in the face of strong inculpatory evidence; thus, in addition to depriving defendants of their 6th Amendment right to a jury verdict, evidentiary weight may be a source of increased sentencing disparities. The current body of scholarship in this area contains virtually no studies that directly explore judicial perceptions of the role of evidentiary weight by speaking to judges candidly about these issues. This is a notable void that limits our ability to understand how and why this influential extralegal factor drives judges to disregard court mandates and sentence based on the strength of the evidence.

This dissertation contributes to the extant literature on extralegal sentencing factors in three ways: first, as indicated above, I identify evidentiary weight as an extralegal sentencing characteristic in trial cases; second, I confirm prior studies which

conclude that forensic evidence leads to longer sentences for convicted defendants (Peterson et al., 1987, 2010, 2013); third, I explore the role of evidentiary weight in trial cases exclusively and show that both evidentiary type and quantity influence sentence severity post conviction; and, finally, through direct dialogue and utilization of hypothetical case scenarios with sentencing judges, I unpack rationales for this influence, concluding that the type and quantity of evidence in a case influences the level of judicial confidence in guilt. Specifically, evidentiary packages perceived by judges to be stronger result in longer sentences for defendants convicted at trial, while weaker cases have the opposite effect. In the pages that follow, I describe how these contributions expand the current body of literature by enhancing our understanding of extralegal evidentiary influences at sentencing. I then conclude with a series of limitations and directions for future research.

Evidentiary Weight Matters

This dissertation systematically establishes that evidentiary weight acts as an extralegal sentencing factor post conviction in trial cases. Results of an initial regression model (see Table 4.3) demonstrate that cases incorporating forensic evidence and cases with more pieces of physical evidence result in longer sentences for defendants convicted of violent crimes in both plea and trial cases. While these findings are informative, they do not effectively demonstrate whether or not judicial decisions drive these disparities. This is because sentencing in plea cases predominantly reflects prosecutorial discretion and evidentiary weight appropriately influences plea offers and negotiations (Bushway, Redlich & Norris, 2014). In contrast, the sentencing phase of trial cases is the purest

reflection of judicial sentencing discretion; it is the arena where judges (as opposed to prosecutors or other court actors) are empowered with the widest discretionary powers (Johnson, 2003). Trials are also the forum in which judges have the greatest opportunity to observe and evaluate the evidence and form strongly grounded opinions regarding evidentiary weight. To better isolate these effects, analyses focused exclusively on trial convictions were performed; these results established that judges impose more punitive sentences when the evidentiary package includes greater quantities of physical evidence and when a portion of the evidence has been forensically examined in a crime laboratory. In contrast, cases with fewer pieces of physical evidence and cases without forensic evidence are associated with shorter sentences for defendants convicted at trial. The presence of eyewitness testimony was not influential. Given that sentencing is a distinct phase that occurs after guilt related issues have been resolved, the influence of type and quantity of physical evidence post conviction in trial cases is counterintuitive. One explanation for these findings is that judicial confidence levels in guilt drive sentencing punitiveness.

To examine this explanation, I engaged state court judges in in-depth qualitative interviews. These data confirmed that judicial perceptions of different evidentiary forms are at the heart of these processes. My interviews with judges revealed that judges consider forensic evidence to be stronger and more objective than non-scientific witness-based evidence (see Chapter 4). As a result, many judges expressed increased confidence in the defendant's guilt when the case contained forensic evidence (see Chapter 5). In accordance with some prior studies (Saks & Faigman, 2008), many judges revere the scientific processes and explained that science is less prone to credibility and reliability

concerns than other evidentiary forms. Yet, even among forensic forms of evidence, judges described a hierarchy of perceived reliability. DNA evidence is considered to be the strongest and most objective form of forensic evidence. Conversely, most judges expressed concern with human error and the subjective aspects of human perceptions. Thus, even among forensic evidentiary forms, those perceived as less subject to human subjective interpretations are considered more reliable. For example, fingerprints are initially analyzed by a computer and subsequently visually examined by experts; this procedure inserts an element of subjectivity into the process. While still perceived as scientific and reasonably reliable, judges generally rank fingerprints lower on the evidentiary scale than DNA.

The hierarchy of perceived strength of forensic evidentiary forms demonstrates that the more evidence is subject to human analyses and judgment, the greater the potential for error, thereby rendering evidentiary perceptions less reliable. Nonetheless, viewed as a group, forensic evidentiary forms were considered to be more objective and reliable than witness evidence. Some judges expressed increased comfort in sentencing the defendant to a longer prison term when the case includes forensic evidence, since such evidence instills confidence in guilt. In contrast, and consistent with the literature (Loftus, 1974; Sanders, 1984; Sherman, Rose & Hoffman, 2011; Skolnock & Shaw, 2001), many judges expressed deep concerns regarding the credibility and reliability of eyewitnesses; public perceptions, media coverage and knowledge of the literature fueled opinions regarding the potential for misidentifications. The potential for human error was at the heart of these concerns.

The evidentiary value assigned to alibi and character witnesses were likewise low compared to forensic evidentiary forms. Yet, in contrast to eyewitnesses, credibility concerns underpinned perceptions of decreased evidentiary value, rather than questions of reliability. Alibi and character witnesses are motivated to help the defendant. The subjective nature of this testimony combined with motive to lie renders uncorroborated alibi evidence as weak. Likewise, absent significant credibility boosts, such as extraordinary high status (e.g. the character witness is the Pope), character witnesses are often presumed to be biased, unhelpful, and subjective. Indeed, their testimony is considered to be the weakest form of evidence among most judges. Thus, viewed as a group, witness-based evidence is viewed as a weaker, more subjective, and less certain form of evidence. Yet to the extent that witness based evidence corroborates other case evidence and reflects a sympathetic “human element,” judges indicated that it can be valuable.

Several factors complicate these findings. First, judges placed value on perceived “objective” evidentiary qualities. Yet even when considering ostensibly objective evidentiary forms, judges tended to focus on subjective evaluations of expert witnesses as opposed to more objective critiques of the scientific processes utilized. Indeed, in discussing forensic expert witnesses, judges focused on showmanship, clarity, and overall testifying abilities (i.e. areas relating to judicial expertise). This finding was consistent among numerous judges and across different forensic evidentiary forms. Thus, while judges expressed an increased confidence in guilt due to their trust in science, many judges failed to mention scientific principles at all and focused instead on subjective presentation skills.

Second, some judges conceded, and others implied, that they lacked an understanding of the scientific processes involved, particularly in DNA cases. Clearly, this lack of knowledge limits the ability of judges to effectively evaluate the reliability of the process utilized by experts. However, these same judges praised DNA evidence for its strength and objective qualities. In contrast, evidentiary analyses that are more easily understood received considerably more criticism (e.g. fingerprints and ballistic identification evidence).

Third, many judges criticized jurors for demanding forensic evidence. Judges spoke of the CSI effect in a critical and judgmental way; jurors demanding forensics were perceived as impediments to effective case processing by overburdening the system with unreasonably high expectations for forensic evidence. This is odd in light of the expressed confidence, certainty, and peace of mind judges associated with forensic evidence in general. Despite the high evidentiary value attributed to forensic evidence, judges were critical of jurors for desiring that same certainty.

Perhaps of greatest concern are judicial perceptions that forensic evidence is objective and provides high certainty in light of the scholarship in this area (Saks & Faigman, 2008; see also Edmond & Roque, 2012). While these results are consistent with prior studies showing that courts place a great deal of confidence in forensic science (Saks & Faigman, 2008), they are not aligned with study results that demonstrate that forensic science, particularly some of the identification sciences, are not nearly as objective and reliable as they are typically perceived (Dror, 2013). Thus, not only is evidentiary weight improperly influencing sentencing, this influence may be based – at least in part – on inaccurate assessments of evidentiary objectivity and value.

Perhaps the most significant contributions of this dissertation are set forth in Chapter 6. A small majority of judges explicitly stated that they consider strength of the evidence at sentencing. Further, these judges acknowledged that their confidence level in a defendant's guilt influences their sentencing decisions. Judges indicated a host of reasons for this influence including the desire to maintain their peace of mind, mitigating perceived guilt phase injustices, hedging against potential errors, accounting for over or under convictions, and avoiding reversal on appeal. Further, the use of vignettes permitted judges to share their reasoning during a sentencing process and demonstrate how these concerns influence sentencing decisions. Through extensive dialogue, judges clearly expressed confidence in guilt ("at least we know he did it" in Scenario B) with cases that included more evidence. Their hypothetical sentences indicated that the inclusion of strong evidence (i.e. Scenario B contains a combination of forensic and eyewitness evidence) resulted in longer average sentences. In contrast, cases with little evidence (i.e. one eyewitness identification in Scenario A) led to uncertainty and doubt, and resulted in shorter sentences on average.

Similar to the quantitative analyses in Chapter 4 and consistent with judicial perceptions of evidentiary type discussed in Chapter 5, evidentiary type was influential. The presence of forensic evidence influenced sentence length. Specifically, in cases that lacked forensic evidence, judges who considered the evidentiary package imposed lower sentences compared to those who did not. These results were corroborated by judicial comments on the strength of forensic evidentiary forms as they reasoned through their sentencing rationale. In contrast, while judges discussed the beneficial effect of eyewitnesses who corroborated the victim's testimony, these acknowledged benefits were

not reflected in the average sentences imposed. Finally, a few judges, who indicated that they do not consider evidentiary weight or confidence in guilt in their sentencing decisions, nevertheless discussed evidentiary factors in describing the processes by which they arrived at their sentencing decisions during the vignette phase of the interview; this suggests that evidentiary weight may be influencing their decisions at least subconsciously.

In the end, this dissertation has systematically established the following:

- Judges perceive forensic evidence to be stronger and more objective than non-scientific witness testimony.
- Greater quantities of evidence have a corroborating effect and increase judicial perceptions of case strength.
- Forensic evidence instills confidence in guilt.
- Confidence level in guilt influences sentence length.
- Perceptions of evidentiary strength influence sentence length.
- Even among those who deny the role of evidentiary weight at sentencing, at least in some cases, part of the rationale for the sentence relies on evidentiary strength. Thus, evidentiary weight is a subconscious influence for certain judges.

These findings are worthy of attention for several reasons. First, the right to a jury trial is a cornerstone of our criminal justice process. History has established that judges vary in their philosophies, goals of punishment, and overall approaches. The possibility that a judge will insert his or her own fact finding observations into the process adds an unconstitutional layer of complexity whereby a defendant may be judged by two fact finders with potentially conflicting conclusions. Yet, can judges be expected to ignore their personal opinions regarding evidentiary strength, when such opinions present moral sentencing dilemmas? These analyses demonstrate that the relationship between

sentencing judge and jury is a complicated one; sentencing judges, who disagree with a jury verdict, may be forced to choose between exceeding their authority and imposing a sentence they perceive to be unjust.

These findings expand the current body of literature in significant ways; yet, they raise critical issues that complicate existing scholarship. First, prior studies have shown that seriousness of the offense and the defendant's criminal history are the most influential factors at sentencing (Bushway & Piehl, 2001; Kramer & Ulmer, 1996). However, the influence of offense severity presupposes that the sentencer believes that the defendant is guilty of the offense. The current findings suggest that the influence of offense seriousness on sentence severity may be conditioned by confidence level in guilt. Since individual judges are moved differently by evidentiary factors, reactions to the same evidentiary package may instill confidence in guilt for some judges while leaving others with grave concerns. Scholars have found that judges utilize substantive rational criteria in their sentencing decisions (Ulmer & Kramer, 1996, p. 384; see also Johnson, 2005; Savelsberg, 1992). According to a focal concerns perspective, judicial decision making is based on three main considerations: blameworthiness of the defendant, protection of the community, and practical concerns (Kramer & Ulmer, 1996; Steffensmeier & Demuth, 2006; Steffensmeier, Ulmer & Kramer, 1998; Ulmer & Johnson, 2004). The current findings add a new perspective to these judicial considerations. Not only does the act upon which the defendant was convicted influence perceptions as to defendant blameworthiness and protection of the community, but, also, the likelihood that the defendant is truly the perpetrator is an influential prerequisite consideration. Indeed, a wrongfully convicted defendant does not pose a threat to the

community nor is he or she blameworthy. Said differently, belief in the correctness of the conviction is necessary before a judge can focus on his or her perceived punishment objectives.

This study further contributes to the current literature by delving into the minds of judges to better understand how these sentencing factors come into play in judicial decision making. The complicated relationship between judge and jury has important sentencing implications. The moral conflict that ensues when a judge is tasked with sentencing a defendant after a perceived questionable verdict has far reaching consequences that necessarily impact other sentencing goals and objectives. Understanding the influence of strength of evidence on sentencing decisions post-verdict, through the words of sentencing judges directly, enhances our understanding not only of judicial perceptions of various evidentiary forms but it also illuminates the complications that arise when the guilt phase fact finder and the sentencer are not aligned in their perspectives. Yet, almost no scholarly focus has been directed toward exploring the impact of judicial doubt in the verdict on sentence severity. While a growing body of literature focuses on expanding our understanding of sentencing in plea cases, including analyses of the shadow of the trial model, little focus is dedicated to trial cases exclusively, and even less is qualitative in nature.

Second, judicial focus on the type and quantity of evidence garnered by law enforcement increase the potential for extralegal influence of sociodemographic considerations but in hidden – and potentially more insidious – ways. Cooney (1994) argues that evidence is socially produced; that is, evidence in criminal cases is dependent on law enforcement priorities and efforts, as well as the ability of victims and defendants

to attract witnesses in support of their cause. Thus, according to Cooney (1994), the status characteristics of the parties to the offense actually *produce* specific types of evidentiary packages. The degree of effort exerted by law enforcement to not only collect multiple types of physical evidence during their investigations, but also expend resources on forensic examination of some or all of that evidence may help to identify certain sentencing disparities often directly attributed to the race and class of victims and offenders. Said differently, part of the explanation for racial and class disparities in sentencing may be the result of their indirect influence through evidentiary weight.

Study Limitations and Directions for Future Research

Several study limitations warrant mention. First, the quantitative data include four different crime types: aggravated assaults, robberies, rapes, and homicides. While crime type is an important component of crime seriousness, these data do not include degrees within crime type. Yet, significant differences may exist in offense severity based on degree. In order to better capture crime severity, injury (measured by whether or not the victim received medical treatment) is combined with crime type. Yet, future studies should include degree information so that this important legally relevant sentencing criteria may be more accurately measured. Since evidentiary strength may influence perceptions of offense severity, more detailed and accurate measures of severity would be particularly helpful in capturing the specific nature and degree of evidentiary influence.

Second, while these quantitative data include numerous forensic evidentiary variables, only a few broad based variables – related to inculpatory evidence – reflect witness based evidence. In contrast, defense witness information (e.g. alibi witnesses,

character witnesses, defense eyewitnesses) are not provided in these analyses. In addition, there is no distinction made in the quantitative data between police and lay witnesses. This information may provide an additional avenue by which to assess the interplay between type of witness and judicial credibility and reliability determinations. Moreover, confession information is not included. Confessions are often the most influential piece of evidence in a case (*Colorado v. Connelly*, 1986) and may further our understanding of the role of evidence at sentencing. Distinctions based on type of witness, direction (prosecution or defense), scope, and quality would also be helpful in identifying distinctions in evidentiary influence that are reflective of the analytical processes engaged by judges.

Third, the quantitative analyses are limited to cases in which the defendant was convicted and sentenced to a period of incarceration. Because the crime types included in these data are very serious, less than 2% of the cases resulted in probationary sentences. Given the small sample size and qualitatively distinct features of probationary sentences, probation cases were excluded from these analyses. Nonetheless, probationary cases may provide an important area of future analysis.

Fourth, the vignettes also did not include any defense information such as defense witnesses, statements, or other physical or documentary evidence. Therefore, the influence of exculpatory evidence on sentence severity could not be directly observed. Further, perceived differences in crime severity between the aggravated assault and robbery vignettes restricted comparisons to analyses within crime type; while these analyses provided valuable data, future construction should better account for differences in perception of crime severity to allow for more direct comparisons.

In light of these study results and limitations, several areas of future research are warranted. First, additional qualitative analyses, exploring the tensions between judges and juries and the ramifications at sentencing when judges disagree with jury verdicts, can enhance our understanding of judicial decision making in trial cases. Second, future studies should explore the connection between evidence production during earlier investigatory case processing phases and sentencing severity. Research directed at illuminating the influence of sociodemographic characteristics of defendants and victims on evidence production and potential connections to sentence severity can enhance our understanding of existing disparities in sentencing based on race and other defendant and victim characteristics. Relatedly, future studies should explore any existing relationship between variations in sentence severity and evidence collection procedures among different jurisdictions; these differences may help to explain jurisdictional variations in sentencing based on evidence collection protocols. Further, qualitative studies with police personnel, regarding evidence collection practices, can provide helpful data to clarify the nature and scope of evidentiary influences. Finally, from a policy perspective, increased police training and monitoring of case processing procedures from early investigatory stages may reduce sociodemographic disparities, as well as evidentiary ones. Increased judicial training and awareness of these influences through education may also help to reduce disparities in some cases.

In the end, these study results suggest that evidentiary weight is an influential extralegal factor at sentencing. Despite its relevance to guilt phase decisions, it can lead to significant disparities when considered during the sentencing stage of case processing. Consciously or not, judges who render sentences based, at least in part, on evidentiary

weight following a jury conviction usurp the role of the jury and interfere with constitutional protections. Despite sincere intentions to mitigate perceived guilt phase injustices, modifying sentences based on confidence level in guilt impedes the imposition of just sentences and ultimately leads to increased sentencing disparities.

References

- ABA Guidelines for Fair Treatment of Crime Victims and Witnesses (1983).
- Albonetti, C. A. (1997). Sentencing under the Federal Sentencing Guidelines: Effects of defendant's characteristics, guilty pleas, and departures on sentence outcomes for drug offenses, 1991-1992. *Law and Society Review*, 31, 789-822.
- Albonetti, C. A. (2011). Judicial discretion in federal sentencing: An intersection of policy priorities and law. *Criminology and Public Policy*, 10, 1151-1155.
- Alleyne v. U.S., 133 S. Ct. 2151 (2013).
- Alvarez, A., & Bachman, R. D. (1996). American Indians and sentencing disparity: An Arizona test. *Journal of Criminal Justice*, 24, 549-561.
- Apprendi v. New Jersey, 530 U.S. 466 (2000).
- Bagaric, M., & Brebner, J. (2002). The solution to the dilemma presented by the guilty plea discount: The qualified guilty plea- I'm pleading guilty only because of the discount! *International Journal of the Sociology of Law*, 30, 51-74.
- Baldus, D. C., Pulaski, C., & Woodworth, G. (1983). Comparative review of death sentences: An empirical study of the Georgia experience. *Journal of Criminal Law and Criminology* 74, 661-753.
- Barter, C., & Renold, E. (2000). 'I wanna tell you a story': Exploring the application of vignettes in qualitative research with children and young people. *International Journal of Social Research Methodology*, 3, 307-323.
- Baskin, D., & Sommers, I. (2010). The influence of forensic evidence on the case outcomes of homicide incidents. *Journal of Criminal Justice*, 38, 1141-1149.

- Baskin, D., & Sommers, I. (2012). The influence of forensic evidence on the case outcomes of assault and robbery incidents. *Criminal Justice Policy*, 23, 186-210.
- Baumer, E. P., Messner, S. F., & Felson, R.B. (2000). The role of victim characteristics in the disposition of murder cases. *Justice Quarterly* 17, 281-307.
- Bennett, M.W. (2014). Confronting cognitive “anchoring effect” and “blind spot” biases in federal sentencing: A modest solution for reforming a fundamental flaw. *Journal of Criminal Law and Criminology*, 104, 489-534.
- Benson, M. L., & Walker, E. (1988). Sentencing the white-collar offender. *American Sociological Review*, 53, 294-302.
- Bibas, S. (2004). Plea bargaining outside the shadow of trial. *Harvard Law Review* 117, 2463-2547.
- Biben, E. (2011). Investigation into the Nassau County police department forensic evidence bureau. Retrieved from <http://www.ig.state.ny.us>.
- Borg, M. J., & Parker, K. F. (2001). Mobilizing law in urban areas: The social structure of homicide clearing rates. *Law and Society Review*, 35, 435-466.
- Briody, M. (2004). The effects of DNA evidence on homicide cases in court. *Australian & New Zealand Journal of Criminology*, 37, 231-252.
- Brodsky, S. L., Griffin, M. P., & Cramer, R. J. (2010). The witness credibility scale: An outcome measure for expert witness research. *Behavioral Sciences and the Law*, 28, 892-907.
- Bushway, S. D., & Forst, B. (2013). Studying discretion in the processes that generate criminal justice sanctions. *Justice Quarterly*, 30, 199-222.

- Bushway, S. D., & Piehl, A. M. (2007). The inextricable link between age and criminal history in sentencing. *Crime and Delinquency*, 53, 156-183.
- Bushway, S. D., & Piehl, A. M. (2001). Judging judicial discretion: Legal factors and racial discrimination in sentencing. *Law and Society Review*, 35, 733-764.
- Bushway, S. D., Redlich, A. D., & Norris, R. J. (2014). An explicit test of plea bargaining in the “shadow of the trial.” *Criminology*, 52, 723-754.
- Cassell, P. G. (2009). In defense of victim impact statements. *Ohio St. J. Crim. L.*, 6, 611-648.
- Clancy, K., Bartolomeo, J., Richardson, D., & Wellford, C. (1981). Sentence decision making: The logic of sentence decisions and the extent and sources of sentencing disparity. *The Journal of Criminal Law and Criminology*, 72, 524-554.
- Clark, S. E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior*, 29, 575-604.
- Clear, T. R. (2008). The effects of high imprisonment rate on communities. *Crime and Justice*, 37, 97-132.
- Cole, S. A. (2009). Forensics without uniqueness, conclusions without individualization: The new epistemology of forensic identification. *Law, Probability and Ri*, 233-255.
- Cole, S. A., & Dioso-Villa, R. (2008). Investigating the CSI effect: Media and litigation crisis in criminal law. *Stan. L. Rev.*, 61, 1335-1374.
- Cole, S. A., & Dioso-Villa, R. (2011). Should judges worry about the “CSI effect”? *Court Review*, 47, 19-31.
- Colorado v. Connelly, 49 U.S. 157 (1986).

- Connolly, D. A., & Gordon, H. M. (2011). He-said-she-said: Contrast effects in credibility assessments and possible threats to fundamental principles of criminal law. *Legal and Criminological Psychology*, 16, 227-241.
- Cooney, M. (1994). Evidence as partisanship. *Law and Society Review*, 28, 833-858.
- Criminal Procedure Law, N.Y.S., §60 (2015).
- Culhane, S. E., & Hosch, H. M. (2004). An alibi witness' influence on mock jurors' verdicts. *Journal of Applied Social Psychology*, 34, 1604-1616.
- Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).
- Demleitner, N. (2005). Constitutional challenges, risk-based analysis and criminal history databases: More demands on the U.S. sentencing commission. *Federal Sentencing Reporter*, 17, 159-165.
- Department of Justice. (2006). *Fact sheet: The impact of United States v. Booker on federal sentencing*. Washington, D.C.
- Devine, D., Buddenbaum, J., Houp, S., Studebaker, N., & Stolle, D. (2009). Strength of evidence, extra evidentiary influence and the liberation hypothesis: Data from the field. *Law and Human Behavior*, 33, 136-148.
- Doerner, J. K. (2012). Gender disparities in sentencing departures: An examination of U.S. federal courts. *Women and Criminal Justice*, 22, 176-205.
- Doerner, J. K., & Demuth, S. (2010). The independent and joint effects of race/ethnicity, gender, and age on sentencing outcomes in U.S. federal courts. *Justice Quarterly*, 27, 1-27.
- Dror, I. (2013). The ambition to be scientific: Human expert performance and objectivity. *Science and Justice*, 53, 81-82.

- Dror, I. E., Champod, C., Langenburg, G., Charlton, D., Hunt, H., & Rosenthal, R. (2011). Cognitive issues in fingerprint analysis: Inter-and intra-expert consistency and the effect of a 'target' comparison. *Forensic Science International*, 208, 10-17.
- Dysart, K. L. (2012). Managing the CSI effect in jurors. *Trial Evidence*, 20.
- Edkins, V. A. (2011). Defense attorney plea recommendations and client race: Does zealous representation apply equally to all? *Law and Human Behavior*, 35, 413-425.
- Edmond, G., & Roque, M. S. (2012-2013). The cool crucible: Forensic science and the frailty of the criminal trial. *Current Issues in Criminal Justice*, 24, 51-68.
- Everett, R. S., & Wojtkiewicz, R. A. (2002). Difference, disparity and race/ethnic bias in federal sentencing. *Journal of Quantitative Criminology*, 18, 189-211.
- Faust, D. & Ziskin, J. (1988). The expert witness in psychology & psychiatry, *Science*, 241, 31-35.
- Federal Sentencing Guidelines, 18 U.S.C. § 3771 (2012).
- Federal Sentencing Guidelines, 18 U.S.C. §§ 3553 et seq. (2012).
- Feeley, M. M. (1979). Perspectives on plea bargaining. *Law & Society Review*, 13, 199-209.
- Feeney, T. (2003). Reaffirming the rule of law in federal sentencing. *Criminal Justice Ethics*, 22, 2-73.
- Feldmeyer, B., & Ulmer, J. (2011). Racial/ethnic threat and federal sentencing. *Journal of Research in Crime and Delinquency*, 48, 238-270.
- Florida Statutes Annotated § 921.141 (2013).

- Forst, B., & Rhodes, W. M. (1982). Structuring the exercise of sentencing discretion in the federal courts. *Federal Probation*, 46, 3-13.
- Forst, B., & Wellford, C. (1981). Punishment and sentencing: Developing sentencing guidelines empirically from principles of punishment. *Rutgers Law Review*, 33, 799-837.
- Frankel, M. E. (1972). *Criminal sentences: Law without order*. N.Y.: Hill and Wang.
- Frankel, M. E. (1972). Lawlessness in sentencing. *Cincinnati Law Review*, 41, 1-54.
- Freiburger, T. L. (2011). The impact of gender, offense type and familial role on the decision to incarcerate. *Social Justice Research*, 24, 143-167.
- Freiburger, T. L., & Hilinski, C. M. (2013). An examination of the interaction of race and gender on sentencing decisions using a trichotomous dependent variable. *Crime and Delinquency*, 59, 59-86.
- Gall v. U.S., 128 S. Ct. 586 (2007).
- Gertner, N. (1999). Circumventing juries, undermining justice: Lessons from criminal trials and sentencing. *Suffolk University Law Review*, 32, 419-439.
- Golding, J. M., Fryman, H. M., Marsil, D. F., & Yozwiak, J. A. (2003). Big girls don't cry: The effect of child witness demeanor on juror decisions in a child sexual abuse trial. *Child Abuse and Neglect*, 27, 1311-1321.
- Goldman, E.F. & Swayze, S. (2012). In depth interviewing with healthcare corporate elites: Strategies for Entry and engagement. *International Journal for Qualitative Methods*, 11 (3), 230-242.
- Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. U.K.: John Wiley and Sons, Ltd.

- Gutheil, T. G. (2003). Types of witnesses. *Focus*, 1, 385-388.
- Hackett, L., Day, A., & Mohr, P. (2008). Expectancy violation and perceptions of rape victim credibility. *Legal and Criminological Psychology*, 13, 323-334.
- Hayes-Smith, R. M., & Levett, L. M. (2011). Jury's still out: How television and crime show viewing influences jurors' evaluations of evidence. *Applied Psychology in Criminal Justice*, 7, 29-46.
- Heinrich, D. P., Thorton, A. E., Morgan, R. M., & Bouhana, N. (2012). An analysis of forensic evidence used in the prosecution of terrorism cases in Britain between 1972 and 2008. *Policing*, 7, 96-108.
- Hills, A. M., & Thomson, D. M. (1999). Should victim impact influence sentences? Understanding the community's justice reasoning. *Behavioral Sciences and the Law*, 17, 661-671.
- Hofer, P. J. (2011). Has Booker restored balance? A look at data on plea bargaining and sentencing. *Federal Sentencing Reporter*, 326-332.
- Hofer, P. J. (2007). United States v. Booker as a natural experiment: Using empirical research to inform the federal sentencing policy debate. *Criminology and Public Policy*, 6, 433-460.
- Hofer, P. J., Blackwell, K. R., & Ruback, R. B. (1999). The effect of the Federal Sentencing Guidelines on inter-judge sentencing disparity. *The Journal of Criminal Law and Criminology*, 90, 239-306.
- Hughes, R. (1998). Considering the vignette technique and its application to the study of drug injecting and HIV risk and safer behaviour. *Sociology of Health and Illness*, 20, 381-400.

- Hughes, T., & Magers, M. (2007). The perceived impact of crime scene investigation shows on the administration of justice. *Journal of Criminal Justice and Popular Culture*, 14, 259-276.
- Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2001). *Criminal interrogations and confessions* (4th ed.). Sudbury, Mass.: Aspen Publications.
- Johnson, B. D. (2003). Racial and ethnic disparities in sentencing departures across modes of conviction. *Criminology*, 41, 449-486.
- Johnson, B. D. (2005). Contextual disparities in guidelines departures: Courtroom social contexts, guidelines compliance and extralegal disparities in criminal sentencing. *Criminology*, 43, 761-796.
- Johnson, B. D., & Betsinger, S. (2009). Punishing the “model minority”: Asian-American criminal sentencing outcomes in federal district courts. *Criminology*, 47, 1045-1090.
- Johnson, B. D., Stewart, E. A., Pickett, J., & Gertz, M. (2011). Ethnic threat and social control: Examining public support for judicial use of ethnicity in punishment. *Criminology*, 49, 401-44.
- Johnson, B. D., & DiPietro, S. M. (2012). The power of diversion: Intermediate sanctions and sentencing disparity under presumptive guidelines. *Criminology*, 50, 811-850.
- Johnson, D., Peterson, J., Sommers, I., & Baskin, D. (2012). Use of forensic science in investigating crimes of sexual violence: Contrasting its theoretical potential with empirical realities. *Violence Against Women*, 18, 193-222.
- Kalven, H., & Zeisel, H. (1966). *The American jury*. Boston, Mass.: Little & Brown.

- Karle, T. W., & Sager, T. (1991). Are the Federal Sentencing Guidelines meeting congressional goals?: An empirical and case law analysis. *Emory Law Journal*, 40, 393-444.
- Kassin, S. M., & Kiechel, K. L. (1996). The social psychology of false confessions: Compliance, internalization, and confabulation. *Psychological Science*, 7, 125-128.
- Kim, Y. S., Barak, G., & Shelton, D. E. (2009). Examining the “CSI effect” in the cases of circumstantial evidence and eyewitness testimony: Multivariate and path analyses. *Journal of Criminal Justice*, 37, 452-460.
- Klaver, J., Lee, Z., & Rose, V. G. (2008). Effects of personality, interrogation techniques and plausibility in an experimental false confession paradigm. *Legal and Criminological Society*, 13, 71-88.
- Kramer, G., Wolbransky, M., & Heilbrun, K. (2007). Plea bargaining recommendations by criminal defense attorneys: Evidence strength, potential sentence and defendant preferences. *Behavioral Sciences and the Law*, 25, 573-585.
- Kramer, J. H., & Ulmer, J. T. (1996, March). Sentencing disparity and departures from guidelines. *Justice Quarterly*, 13, 81-106.
- Lear, E. T. (1993). Is conviction irrelevant? *U.C.L.A. Law Review*, 40, 1179-1239.
- Leippe, M. R., Eisenstadt, D., & Rauch, S. M. (2009). Cueing confidence in eyewitness identifications: Influence of biased lineup instructions and pre-identification memory feedback under varying lineup conditions. *Law and Human Behavior*, 33, 194-212.

- Levett, L. M. (2013). Co-witness information influences whether a witness is likely to choose from a lineup. *Legal and Criminological Psychology*, 18, 168-180.
- Lichtblau, C. (2004). Report questions the reliability of an F.B.I. ballistics test. *New York Times*.
- Lockett v. Ohio, 438 U.S. 586 (1978).
- Lofland, J., Snow, D., Anderson, L., & Lofland, L. H. (2006). *Analyzing social settings: A guide to qualitative observation and analysis* (4th ed.). Belmont, Cal.: Wadsworth.
- Loftus, E. F. (1974). Reconstructing memory: The incredible eyewitnesses. *Psychology Today*, 8, 116-119.
- Lynch, G. E. (2007-2008). *Ohio St. J. Crim. L.*, 5.
- Lynch, G. E. (2009). Marvin Frankel: A reformer reassessed. *Federal Sentencing Reporter*, 21, 235-241.
- Mah, C. L., Taylor, E., Hoang, S., & Cook, B. (2014). Using Vignettes to Tap Into Moral Reasoning in Public Health Policy: Practical Advice and Design Principles From a Study on Food Advertising to Children. *American Journal of Public Health*, 104, 1826-1832.
- Maxwell, J. A. (2005). *Qualitative research design: An interactive approach* (2nd ed.). Thousand Oaks, Cal.: Sage Publications, Inc.
- McMillan et.al. v. Pennsylvania, 477 U.S. 79 (1986).
- McQuiston-Surrett, D., & Saks, M. J. (2009). The testimony of forensic identification science: What expert witnesses say and what factfinders hear. *Law and Human Behavior*, 33, 436-453.

- Mustard, D. B. (2001, April). Racial, ethnic, and gender disparities in sentencing: Evidence from the U.S. federal courts. *Journal of Law and Economy*, 44, 285-315.
- Nadler, J., & Rose, M. R. (2003). Victim impact testimony and the psychology of punishment. *Cornell Law Review*, 88, 419-454.
- Newman, J. O. (2002). Remembering Marvin Frankel: Sentencing reform but not these guidelines. *Federal Sentencing Reporter*, 14, 319-321.
- Nir, E., & Griffiths, E. (2016). Sentencing on the evidence. *Criminal Justice Policy Review*, forthcoming.
- Nowacki, J. S. (2015). Race, Ethnicity and Judicial Discretion: The Influence of the United States v. Booker Decision. *Crime & Delinquency*, 61, 1360-1385.
- Osborne, Y. H., & Rappaport, N. B. (1985). Sentencing severity with mock jurors: Predictive validity of three variable categories. *Behavioral Sciences and the Law*, 3, 467-473.
- Paternoster, R. (2011). Racial disparity under the federal sentencing guidelines pre- and post-Booker: Lessons not learned from research on the death penalty. *American Society of Criminology*, 10, 1063-1072.
- Paternoster, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology*, 36, 859-866.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, Cal.: Sage Publications, Inc.
- Penal Law, N.Y.S. § 10(3) (2012).
- Penal Law, N.Y.S. § 110.00 (2012).

Penal Law, N.Y.S. § 120.00 (2012).

Penal Law, N.Y.S. § 160.15 (2012).

Penal Law, N.Y.S. § 70 (2012).

Peterson, J. L., Hickman, M. J., Strom, K. J., & Johnson, D. J. (2013). Effect of forensic evidence on criminal justice case processing. *Journal of Forensic Science*, 58, S78-S90.

Peterson, J. L., Mihajlovic, S., & Gilliland, M. (1984). Forensic evidence and the police-The effects of scientific evidence on criminal investigations. *National Institute of Justice*, 1-234.

Peterson, J. L., Ryan, J. P., Houlden, P. J., & Mihajlovic, S. (1987). The uses and effects of forensic sciences in the adjudication of felony cases. *Journal of Forensic Sciences*, 32, 1730-1754.

Peterson, J. L., Sommers, I., Baskin, D., & Johnson, D. (2010). The role and impact of forensic evidence in the criminal justice process. *U.S. Department of Justice*, 1-143.

Phillips, S. W. (2008). Mandatory arrest: Loosely coupled organisations, situation variables, and the arrest decision. *International Journal of Police Science and Management*, 10, 374-387.

Porter, S., & Brinke, L. T. (2009). Dangerous decisions: A theoretical framework for understanding how judges assess credibility in the courtroom. *Legal and Criminological Psychology*, 14, 119-134.

- Ramchand, R., MacDonald, J. M., Haviland, A., & Morral, A. R. (2009). A developmental approach for measuring the severity of crimes. *Journal of Quantitative Criminology*, 25, 129-153.
- Roberts, A., & Lyons, C. J. (2009). Victim-offender racial dyads and clearance of lethal and nonlethal assault. *Journal of Research in Crime and Delinquency*, 46, 301-326.
- Roberts, J. V. (1994). The role of criminal records in the federal sentencing guidelines. *Criminal Justice Ethics*, 13, 21-30.
- Robinson, P. H. (2008). *Distributive principles of criminal law: Who should be punished how much*. New York, N.Y.: Oxford University Press.
- Rodriquez, S. F., Curry, T. R., & Lee, G. (2006). Gender differences in criminal sentencing: Do effects vary across violent, property, and drug offenses? *Social Science Quarterly*, 87, 318-339.
- Saks, M. J., & Faigman, D. L. (2008). Failed forensics: How forensic science lost its way and how it might yet find it. *Annual Review of Law and Social Science*, 4, 149-171.
- Saks, M. J., & Koehler, J. J. (2005). The coming paradigm shift in forensic identification science. *Science*, 309, 892-895.
- Samuel, W., & Moulds, E. (1986). The effect of crime severity on perceptions of fair punishment: A California case study. *Journal of Criminal law and Criminology*, 77, 931-948.
- Sanders, R. (1984). Helping the jury evaluate eyewitness testimony: The need for additional safeguards. *American Journal of Criminal Law*, 12, 189-220.

- Sandys, M., Pruss, H. C., & Walsh, S. M. (2009). Aggravations and mitigations: Findings and implications. *The Journal of Psychiatry and Law*, 37, 189-235.
- Savelsberg, J. T. (1992). Law that does not fit society: Sentencing guidelines as a neoclassical reaction to the dilemmas of substantivized law. *American Journal of Sociology*, 97, 1346-1381.
- Scott, R. W. (2011). Race disparity under advisory guidelines. *Criminology and Public Policy*, 10, 1129-1138.
- Sharps, M. J., Hess, A. B., Casner, H., Ranes, B., & Jones, J. (2007). Eyewitness memory in context: Toward a systematic understanding of eyewitness evidence. *Forensic Examiner*, 16, 20-27.
- Shelton, D. E., Barak, G., & Kim, Y. S. (2011). Studying juror expectations for scientific evidence: A new model for looking at the CSI myth. *Court Review*, 47, 8-18.
- Shermer, L. O., Rose, K. C., & Hoffman, A. (2011). Perceptions and credibility: Understanding the nuances of eyewitness testimony. *Journal of Contemporary Criminal Justice*, 27, 183-203.
- Skolnick, P., & Shaw, J. I. (2001). A comparison of eyewitness and physical evidence on mock-juror decision making. *Criminal Justice and Behavior*, 28, 614-630.
- Smith, S. M., Stinson, V., & Patry, M. W. (2011). Fact of Fiction? The myth and reality of the CSI effect. *Court Review*, 47, 4-7.
- Snortum, J. R., & Riva, P. R. (1990). Police documentation of drunk-driving arrests: Jury verdicts and guilty pleas as a function of quantity and quality evidence. *Journal of Criminal Justice*, 18, 99-116.

- Spohn, C. (2000). *Offender race and case outcomes: Do crime seriousness and strength of evidence matter? (NCJRS/184774)*. Retrieved from <http://www.ncjrs.gov/pdffiles1/nij/grants/184774.pdf>
- Spohn, C. (2011). Unwarranted disparity in the wake of Booker/Fanfan decision: Implications for research and policy. *Criminology and Public Policy*, 10, 1119-1127.
- Spohn, C., & Cederblom, J. (1991). Race and disparities in sentencing: A test of the liberation hypothesis. *Justice Quarterly*, 8, 305-327.
- Spohn, C., & Holleran, D. (2000). The imprisonment penalty paid by young, unemployed black and Hispanic male offenders. *Criminology*, 38, 281-306.
- Starr, S. (2013). Did Booker increase disparity? Why the evidence is unpersuasive. *Federal Sentencing Reporter*, 25, 323-326.
- Steffensmeier, D., & Demuth, S. (2000). Ethnicity and sentencing outcomes in U.S. federal courts: Who is punished more harshly? *American Sociological Review*, 65, 705-729.
- Steffensmeier, D., & Demuth, S. (2006). Does gender modify the effects of race-ethnicity on criminal sanctioning? Sentences for male and female white, black, and Hispanic defendants. *Journal of Quantitative Criminology*, 22, 241-261.
- Steffensmeier, D., Ulmer, J., & Kramer, J. (1998). The interaction of race, gender, and age in criminal sentencing: The punishment cost of being young, black and male. *Criminology*, 36, 763-797.
- Tiede, L. B. (2009). The impact of the federal sentencing guidelines and reform: A comparative analysis. *Justice System Journal*, 30, 34-49.

U.S. v. Booker, 543 U.S. 220 (2005).

U.S. v. Fanfan, 542 U.S. 296 (2004).

U.S. v. Gaudin, 515 U.S. 506 (1995).

U.S. v. Juarez-Ortega, 866 F.2d 747 (5th Cir. 1989)

Ulmer, J., & Johnson, B. (2004). Sentencing in context. *Criminology*, 42, 137-178.

Ulmer, J., Light, M. T., & Kramer, J. (2011). The “liberation” of federal judges’ discretion in the wake of the Booker/Fanfan decision: Is there increased disparity and divergence between courts? *Justice Quarterly*, 28, 799-837.

Ulmer, J. T., Eisenstein, J., & Johnson, B. D. (2010). Trial penalties in federal sentencing: Extra-guidelines factors and district variations. *Justice Quarterly*, 27, 560-592.

Ulmer, J. T., & Kramer, J. H. (1996). Court communities under sentencing guidelines: dilemmas of formal rationality and sentencing disparity. *Criminology*, 34, 383-408.

United States Sentencing Commission. (1991). *The federal sentencing guidelines: A report on the operation of the guidelines system and short-term impacts on disparity in sentencing, use of incarceration, and prosecutorial discretion and plea bargaining* (137987). Washington, DC: Government Printing Office.

United States Sentencing Commission. (2012). *Report on the continuing impact of United States v. Booker on federal sentencing*. Washington D.C.: Government Printing Office.

von Hirsch, A. (1992). Proportionality in the philosophy of punishment. *Crime and Justice*, 16, 55-98.

- Walsh, A. (1990). Standing trial versus coping a plea: Is there a penalty? *Journal of Contemporary Criminal Justice*, 6, 226-236.
- Wells, G. L., & Quinlivan, D. S. (2008). 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.
- Wessel, E., Drevland, G. C., Eilertsen, D. E., & Magnussen, S. (2006). Credibility of the emotional witness: A study of ratings by court judges. *Law Human Behavior*, 30, 221-230.
- Wheeler, S., Weisburd, D., & Body, N. (1982). Sentencing the white collar offender: Rhetoric and reality. *American Sociological Review*, 47, 641-659.
- Wojcikiewicz, J. (2013). Judges' attitude towards scientific evidence. *Revija za Kriminalistiko in Kriminologijo*, 64, 249-255.
- Wolfgang, M. E. (1985). *National survey of crime severity*. Washington D.C.: U.S. Department of Justice.
- Wu, J., & D'Angelo, J. M. (2014). Unwarranted disparity in federal sentencing: Noncitizen crime as a social/group threat. *Criminal Justice Review*, 39, 58-80.
- Wu, J., & Spohn, C. (2009). Does an offender's age have an effect on sentence length?: A meta-analytic review. *Criminal Justice Policy Review*, 20, 379-413.
- www.ats.ucla.edu/stat/stata/dae/ztnb/htm.
- www.icpsr.umich.edu/icpsr/ICPSR/studies/29203/documentation.
- www.innocenceproject.org/understand/Eyewitness-Misidentification.php

www.nccourts.org/Courts/CRS/Councils/spac/Sentencing/Punishment.asp.

www.revisor.mn.gov/court-rules/rule.php.

www.ussc.gov/Guidelines/2011_Guidelines/Manual_PDF/Index.cfm

Table 3.1: Judicial Characteristics

Judge	Years on Bench	Gender	Background	Location	Elected or Appointed?
1	16 years	Male	Prosecutor	Suburban	Elected
2	28 years	Male	Defense	Urban	Appointed
3	25 years	Female	Both	Urban	Appointed
4	18 months	Female	Prosecutor	Urban	Elected
5	9 years	Male	Prosecutor	Urban	Elected
6	12 years	Female	Law secretary	Urban	Elected
7	18 years	Male	Prosecutor	Urban	Elected
8	17 years	Male	Law Secretary	Urban	Appointed
9	27 years	Male	Prosecutor	Urban	Elected
10	9 years	Male	Prosecutor	Suburban	Elected
11	11 years	Male	Defense	Urban	Appointed
12	14 years	Male	Both	Suburban	Elected
13	27 years	Male	Prosecutor	Urban	Appointed
14	14 years	Male	Prosecutor	Urban	Appointed
15	3 years	Female	Defense	Suburban	Elected
16	9 years	Male	Both	Suburban	Elected
17	24 years	Male	Prosecutor	Urban	Appointed
18	19 years	Male	Both	Suburban	Appointed
19	16 years	Male	Both	Urban	Elected
20	19 years	Male	Defense	Rural	Appointed
21	25 years	Male	Defense	Suburban	Appointed
22	26 years	Male	Prosecutor	Suburban	Elected
23	33 years	Male	Defense	Urban	Elected
24	28 years	Male	Prosecutor	Urban	Appointed
25	9 years	Male	Prosecutor	Urban	Appointed
26	28 years	Male	Defense	Urban	Appointed
27	6 years	Male	Prosecution	Urban	Elected
28	17 years	Female	Both	Urban	Elected
29	18 years	Male	Defense	Urban	Elected
30	12 years	Female	Both	Rural	Elected
31	18 years	Male	Prosecutor	Rural	Elected
32	21 years	Male	Both	Rural	Appointed
33	20 years	Male	Defense	Rural	Appointed
34	7 years	Male	Both	Rural	Elected
35	20 years	Male	Prosecutor	Rural	Elected
36	19 years	Male	Both	Rural	Elected
37	25 years	Male	Defense	Rural	Elected
38	17 years	Female	Both	Urban	Elected
39	30 years	Male	Prosecution	Rural	Elected
40	2 years	Male	Both	Rural	Elected
41	23 years	Male	Prosecutor	Urban	Elected

Table 4.1. Descriptive Statistics (N=484)

	MEAN/FREQUENCY	SD/%	RANGE
DEPENDENT VARIABLE			
Sentence length in months [†]	166.57	300.28	1-2760
INDEPENDENT VARIABLES			
Criminal History			
Number of prior convictions	5.20	6.61	0-78
Mode of Conviction			
Plea	326	67.4	
Trial*	158	32.6	
Offense Seriousness			
Homicide*	126	26.0	
Agg. assault with treatment	36	7.4	
Agg. assault no treatment	132	27.3	
Robbery with treatment	14	2.9	
Robbery no treatment	112	23.1	
Rape with treatment	57	11.8	
Rape no treatment	7	1.4	
Quantity of Evidence	1.19	1.75	0-8
Presence of a Forensic Lab Report	203	41.9	
Presence of Eyewitness(es)	278	57.4	
CONTROL VARIABLES			
Male Defendant	448	92.6	
Defendant Race			
African American	233	48.1	
Latino	61	12.6	
Other (White, Asian, other,	190	39.2	
Male Victim	296	61.2	
Victim Race			
African American	161	33.3	
Latino	65	13.4	
Other (White, Asian, other,	258	53.3	
Site			
Los Angeles	185	38.2	
Indiana*	299	61.8	

[†]Untransformed

* Reference category

ABBREVIATION: SD = standard deviation.

Table 4.2. Evidentiary Weight by Crime Type (N=484)

	HOMICIDE (N=126)	RAPE (N=64)	ROBBERY (N=126)	AGGRAVATED ASSAULT (N=168)
Quantity of Evidence^{a,b}				
Mean	3.135	1.250	0.318	0.363
SD	1.978	1.447	0.677	0.836
Range	0-8	0-6	0-4	0-4
Presence of a Forensic Lab Report^{a,b}				
Percent of cases	86.5%	48.4%	25.4%	18.5%
Presence of Eyewitness(es)^{a,c}				
Percent of cases	72.2%	23.4%	47.6%	66.7%

ABBREVIATION: SD = standard deviation.

^a Levene statistic in homogeneity of variance test is significant and thus equal variances cannot be assumed; Games Howell post-hoc tests are conducted

^b All mean differences are significant except the contrast between Aggravated Assault and Robbery

^c All mean differences are significant except the contrast between Aggravated Assault and Homicide

Table 4.3. Ordinary Least Squares Regression Predicting Log of Sentence Length for Defendants Convicted of Violent Crimes (N=483)

	Model 1			Model 2		
	B	SE	% change	B	SE	% change
Number of prior convictions	.016	.010	1.61	.013	.010	1.31
Plea ^a	-.571***	.121	-43.50	-.605***	.120	-45.39
Offense Seriousness ^b						
Agg. assault with treatment	-2.028***	.224	-86.84	-1.582***	.243	-79.44
Agg. Assault no treatment	-2.984***	.152	-94.94	-2.433***	.188	-91.22
Rape with treatment	-1.065***	.185	-65.53	-.728**	.231	-51.71
Rape no treatment	-1.500**	.439	-77.69	-1.093*	.456	-66.47
Robbery with treatment	-2.122***	.318	-88.02	-1.375***	.340	-74.72
Robbery no treatment	-1.756***	.149	-82.73	-1.205***	.197	-70.03
Defendant Race ^c						
African American				.007	.140	.702
Hispanic				-.000	.204	.00
Male Defendant				-.228	.192	-20.39
Victim Race ^c						
African American				-.113	.141	-10.68
Hispanic				-.152	.195	-14.10
Male Victim				.056	.124	5.76
Site: Los Angeles ^d				-.441**	.130	-35.66
Quantity of Evidence				.118*	.046	12.52
Forensic Lab Report				.415**	.144	51.44
Eyewitness(es)				-.074	.113	-7.13
Intercept	5.759	.121		5.546	.299	
Adjusted R ²	.544			.582		
F	73.142***			38.418***		
St. Error of the Estimate	1.127			1.079		

ABBREVIATIONS: b = coefficient; SE = standard error.

^a Ref. category is trial; ^b Ref. category is homicide; ^c Ref. category is other; ^d Ref. category is Indiana

*p<.05; **p<.01; ***p<.001

Table 4.4. Ordinary Least Squares Regression Predicting Log of Sentence Length for Defendants Convicted of Violent Crimes at Trial (N=157)

	B	SE	% change
Number of prior convictions	.033*	.017	3.36
Offense Seriousness ^a			
Agg. assault with treatment	-3.791***	1.084	-97.74
Agg. assault no treatment	-2.958***	.374	-94.81
Rape with treatment	-.694	.440	-50.04
Rape no treatment	-1.505*	.707	-77.80
Robbery with treatment	-1.251**	.460	-71.38
Robbery no treatment	-1.092**	.344	-66.45
Defendant Race ^b			
African American	.100	.243	10.51
Hispanic	.001	.293	.100
Male Defendant	.202	.352	22.38
Victim Race ^b			
African American	-.182	.240	-16.64
Hispanic	.221	.259	24.73
Male Victim	.154	.226	16.65
Site: Los Angeles ^c	-.596*	.233	-44.90
Quantity of Evidence	.150*	.063	16.18
Forensic Lab Report	.600*	.261	82.21
Eyewitness(es)	-.305	.195	-26.29
Intercept	4.960	.548	
Adjusted R ²	.644		
F	17.726***		
St. Error of the Estimate	.99769		

ABBREVIATIONS: b = coefficient; SE = standard error.

^a Ref. category is homicide; ^b Ref. category is other; ^c Ref. category is Indiana

*p<.05; **p<.01; ***p<.001

Table 4.5. Ordinary Least Squares Regression Predicting Log of Sentence Length for Defendants Convicted of Violent Felonies via Pleas (N=325)

	B	SE	% change
Number of prior convictions	.010	.011	1.01
Offense Seriousness ^a			
Agg. assault with treatment	-1.471***	.279	-77.03
Agg. assault no treatment	-2.316***	.230	-90.13
Rape with treatment	-.750**	.283	-52.76
Rape no treatment	-.761	.608	-53.28
Robbery with treatment	-1.225*	.545	-70.62
Robbery no treatment	-1.094***	.259	-66.51
Defendant Race ^b			
African American	-.047	.175	-4.59
Latino	.148	.292	15.95
Male Defendant	-.466*	.232	-37.25
Victim Race ^b			
African American	-.013	.175	-1.29
Latino	-.534	.290	-41.37
Male Victim	-.096	.154	-9.15
Site: Los Angeles ^c	-.311	.173	-26.73
Quantity of Evidence	.076	.068	7.90
Forensic Lab Report	.429*	.185	53.57
Eyewitness(es)	-.007	.144	-.70
Intercept	5.148	.336	
Adjusted R ²	.448		
F	16.542***		
St. Error of the Estimate	1.09469		

ABBREVIATIONS: b = coefficient; SE = standard error.

^a Ref. category is homicide; ^b Ref. category is other (White, Asian, other, unknown); ^c Ref. category is Indiana

*p<.05; **p<.01; ***p<.001

Table 6.1 Description of Vignettes

Scenario	Forensic Evidence	Witnesses	Crime Type	Victim Injury	Identification	Circumstance
A	No	No	Agg. Assault	Stiches	Victim only	Bar fight
B	Yes	Yes	Agg. Assault	Stiches	Witnesses	Dispute at party
C	No	Yes	Robbery	Stiches	Witnesses	Cell phone robbery on street
D	Yes	No	Robbery	Stitches	Victim only	Robbery of necklace on street

Table 6.2 Sentences Imposed by Judges in Vignettes

JUDGE	SCENARIO A AGG. ASSAULT <i>No forensic</i> <i>No witnesses</i>	SCENARIO B AGG. ASSAULT <i>Forensic</i> <i>Witnesses</i>	SCENARIO C ROBBERY <i>No forensic</i> <i>Witnesses</i>	SCENARIO D ROBBERY <i>Forensic</i> <i>No witnesses</i>	DIRECT INQUIRY: <i>CONSIDER</i> <i>EVIDENTIARY</i> <i>STRENGTH?</i>
1	3.5 years	4 years	5 years	7 years	Yes
2	5 years	5 years*	Less than 5**	5 years	Other
3	Probation	Alternative	5 years	5 years	No
4	1 year	Less than 5	5 years	5 years	Other
5	2-3 years	5 years	5 years	7-8 years	No
6	3 years	6 years	5 years	10 years	Yes
7	5 years*	3 years	5-7 years	10-15 years	No
8	3 years	7 years	5-7 years	8 years	Yes
9	3 years	5 years	6-7 years	10 years	Yes
10	Split	1 year	5 years	10 years	Yes
11	1 year	1 year	5 years	5 years	Other
12	5 years or less	7 years	10 or less	15 years	Yes
13	10 years or less	5 years	8-10 years	5 years	No
14	5 years	Split	5 years	5 years	Yes
15	5 years	4 years	Split	7-8 years	Yes
16	2-3 years	8 years	6 years	11 years	Yes
17	9 years	Split	9 years	10-15 years	No
18	7 years	5 years*	15 years	5 years or less	Other
19	Probation	8 years	5 years	12 years	No
20	5 years or less	9 years	10 years	6 years	Yes
21	1 year	6 years	5 years	7 years	Yes
AVERAGE***	3.7	4.5	6.5	8.4	
AVERAGE YES	2.8	5.1	5.8	8.4	
AVERAGE NO	5.0	3.4	7.6	8.5	
RANGE	Probation to 10 yrs or less	Alternative to 9 years	5 – 15 years	5 – 15 years	

* Plus post-supervision release ** Minimum stay and lengthy post-supervision release

***Means calculated on most punitive sentence when range provided, probation coded as 0, and split coded as .5. Split = 6 months incarceration and 5 years of probation.

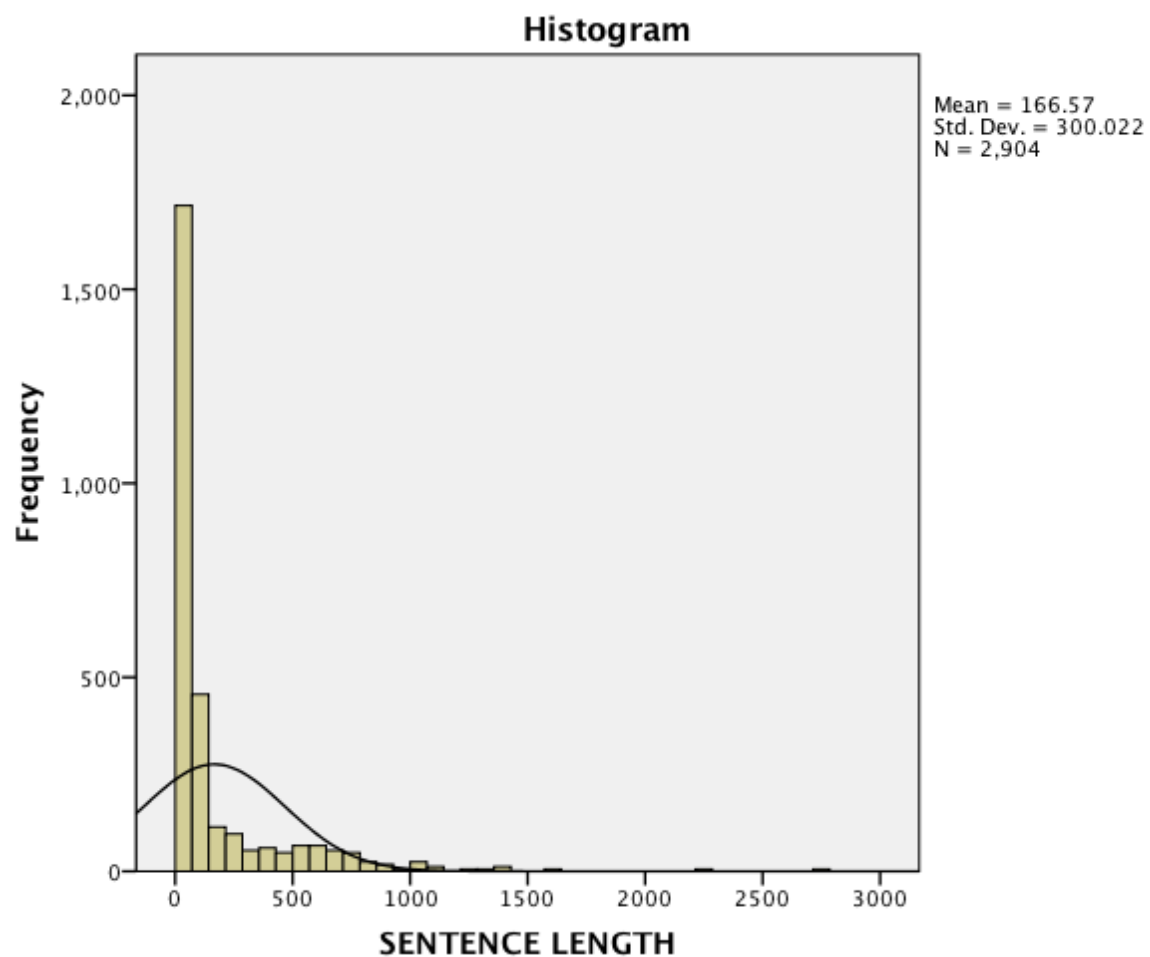
Alternative = alternative program to incarceration

Average Yes = mean sentence for judges who mentioned evidence in sentencing rationale

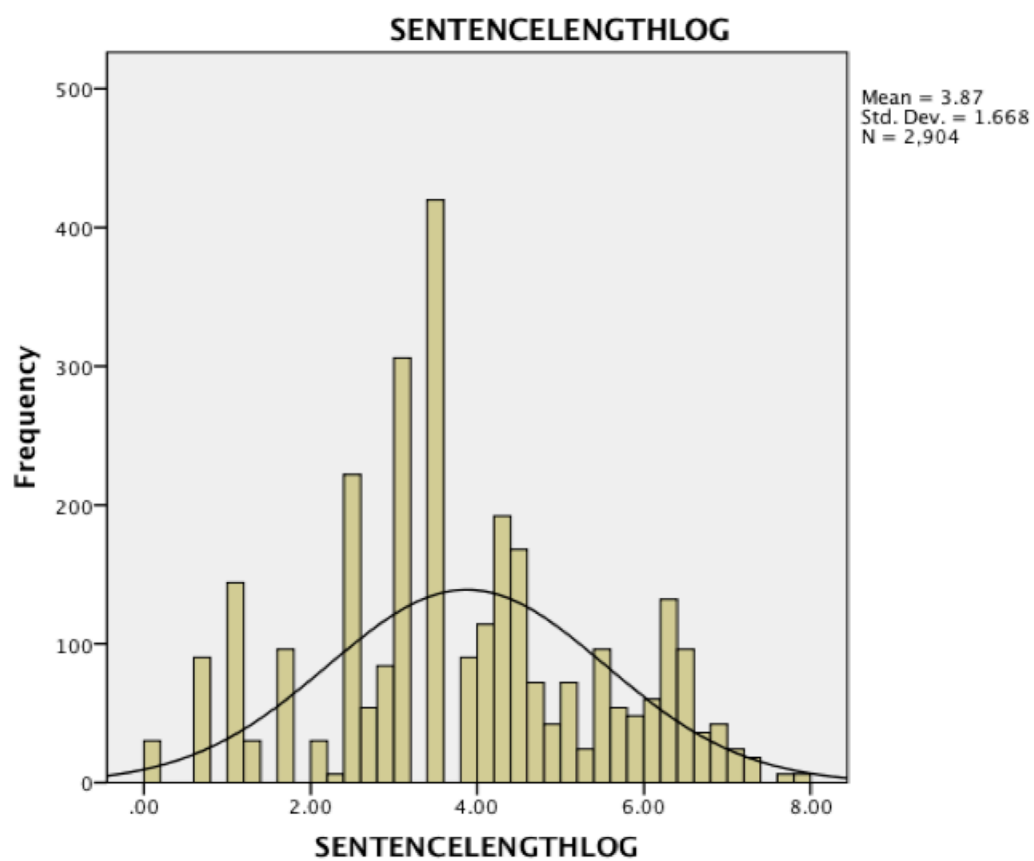
Average No = mean sentence for judges who did not mention evidence in sentencing rationale

BOLD = most punitive within judge Shaded = judge considered evidence

Graph 4.1: Distribution of Sentence Length



Graph 4.2: Distribution of Logged Sentence Length



APPENDIX 3.1

Principal Investigator: Esther Nir

Project Title: Judicial Considerations in Sentencing Determinations

INFORMED CONSENT FORM

1. Introduction

You are invited to participate in a research study that is being conducted by Esther Nir, a doctoral student in the Criminal Justice Department at Rutgers University. This form is designed to provide you with all of the information you will need before you decide whether to participate in this study or to decline participation. It is entirely your choice. If you decide to take part in this study, you can change your mind later on and withdraw from the research study at any time, without penalty. This decision will not adversely affect your relationship with the researcher or with Rutgers University. You have been chosen to participate in this study because you have experience in sentencing defendants who have been convicted of criminal offenses. I plan to interview approximately 50 New York State Supreme Court judges (criminal division). Each interview is expected to last approximately 60-75 minutes.

2. What is the purpose of this research?

The purpose of this research is to study the factors that judges consider in reaching sentencing decisions, as well as judicial thoughts about sentencing practices in general. These interviews are designed to gather information regarding the thought processes and rationales that guide judicial sentencing decisions. This information will help advance the general body of literature on sentencing practices.

3. What procedures are involved?

I will be conducting 60-75 minute interviews with judges regarding the factors they consider in sentencing criminal defendants. You can expect to be asked questions about the factors you consider in sentencing criminal defendants, your thoughts about the jury system and individualized sentencing as well as your views on general sentencing practices. I will also present you with several hypothetical case scenarios and ask you questions regarding how you would handle those situations. You will also be asked basic demographic questions.

This informed consent form was approved by the Rutgers University Institutional Review Board for the Protection of Human Subjects on 5/26/2014; approval of this form expires on 5/25/2015.

With your permission, the interview will be conducted in your chambers/office. Alternatively, the interview can take place at a location in which you feel most comfortable. I would like to audiotape the interview. This will assist me in accurately recording the information that you provide during our interview. The audio recording will later be transcribed. Following transcription, the recording will be erased. If you prefer that this interview not be recorded, notes will be taken instead. There will be no future contact with me once the conversation is completed and you will face no future obligations. Although I ask that you sign your name on the consent form, I will not include your name on the tape recording or on your interview transcript so there will be no information that can directly identify you. The consent forms will be stored in a separate, locked file.

Initial _____

4. Potential risks

There are no foreseeable risks or discomforts associated with your participation in this study. In the event that you find a particular question asked to be uncomfortable, you are not required to answer any question that you prefer not to answer and you may end the interview at any time. Serious measures will be taken to ensure that no breach of confidentiality occurs. There will be no information directly linking you with any information revealed in the interview. Only basic demographic questions will be asked and I will not ask for any personal information that might identify you. The audiotape of the interview will be destroyed after the interview is transcribed and any inadvertent identifiers will be removed during transcription.

5. Benefits

This study is not designed to benefit you directly. Rather, the information gathered will help advance the general body of literature on judicial sentencing practices by focusing on the thoughts, opinions and perspectives of the actual decision makers. Ultimately, through publications resulting from this project, this work may provide an information-sharing tool that can be used to educate new judges as to sentencing procedures as well as inform other criminal justice actors of current sentencing practices.

This informed consent form was approved by the Rutgers University Institutional Review Board for the Protection of Human Subjects on 5/26/2014; approval of this form expires on 5/25/2015.

6. Voluntary participation and withdrawal

Participation in this research is voluntary. You do not have to be in this study. If you decide to be in this study and change your mind, you have the right to drop out at any time. You may skip questions or stop the interview at any time without any penalty.

7. Confidentiality

This research is confidential. Confidential means that the research records will include some information about you and this information will be stored in such a manner that some linkage between your identity and the response in the research exists. The only personal information collected about you includes your gender, ethnicity, age and education. Please note that we will keep this information confidential by limiting individual's access to the research data and keeping it in a secure location. Further, there will be no information directly linking you with any information revealed in the interview. The interview data, excluding identifiers, will be saved on a password-protected computer. After interview transcription, the recording of this interview will be destroyed. A study number rather than your name will be used on interview transcripts and other study materials (excluding this consent form which will be stored in a locked file, separate from other study materials). Your name and other facts that might point to you will not appear when we present this study or publish its results. The only individuals able to view your identifiable information will be study researchers.

The research team and the Institutional Review Board at Rutgers University are the ones that will be allowed to see the data, except as may be required by law. If a report of this study is published, or the results are presented at a professional conference, only group results will be stated. All study data will be kept for 3 years. Upon request, study results will be provided to you at the conclusion of this study.

Initial _____

8. Contact Person:

If you have any questions about this study contact Esther Nir at 201-388-7078 or at enir@rutgers.edu or Dr. Elizabeth Griffiths at 973-353-3303 or at elizabeth.griffiths@rutgers.edu

This informed consent form was approved by the Rutgers University Institutional Review Board for the Protection of Human Subjects on 5/26/2014; approval of this form expires on 5/25/2015.

If you have any questions about your rights as a research subject, you may contact the IRB Administrator at Rutgers University at:

Rutgers University, the State University of New Jersey
Institutional Review Board for the Protection of Human Subjects
Office of Research and Sponsored Programs
3 Rutgers Plaza
New Brunswick, NJ 08901-8559
Tel: 848-932-0150
Email: humansubjects@orsp.rutgers.edu

You will be given a copy of this consent form for your records.

Sign below if you agree to participate in this research study:

Subject (Print) _____

Subject Signature _____ Date _____

Principal Investigator Signature _____ Date _____

INITIAL _____

This informed consent form was approved by the Rutgers University Institutional Review Board for the Protection of Human Subjects on 5/26/2014; approval of this form expires on 5/25/2015.

AUDIO/VIDEOTAPE ADDENDUM TO CONSENT FORM

You have already agreed to participate in a research study entitled **Judicial Considerations in Sentencing Determinations** conducted by **Esther Nir**. We are asking for your permission to allow us to **audiotape** our interview as part of the research study. You do not have to agree to be recorded in order to participate in the main part of the study.

The recording will assist us in accurately documenting the information that you provide during our interview. The recording will be stored in a locked file cabinet with no link to your identity until transcription. After the interview is transcribed, the audio recording will be destroyed. We will not include your name on the tape recording or on your interview transcript so that there will be no information that can directly identify you. The transcript of the audio-recording will be filed by number with all identifiers removed from the transcript.

Your signature on this form grants the investigator named above permission to record you as described above during participation in the above-referenced study. The investigator will not use the recording(s) for any other reason than that/those stated in the consent form without your written permission.

Subject (Print) _____

Subject	Signature	_____	Date

Principal Investigator Signature _____ Date _____

This informed consent form was approved by the Rutgers University Institutional Review Board for the Protection of Human Subjects on 5/26/2014; approval of this form expires on 5/25/2015.

APPENDIX 3.2

INTERVIEW GUIDE NEW YORK STATE SUPREME COURT JUDGES

Thank you for agreeing to meet with me. I would like to start our conversation with some basic background questions.

INTERVIEWEE BACKGROUND:

1. How long have you been on the bench?
2. How long have you been handling criminal cases on the bench?
3. Have you ever been assigned – in your capacity as a judge – to handle civil cases?

IF YES: How would you describe the types of civil cases that you handled? When (and for how long) did you handle those cases?

4. How would you describe the crime types that you have managed throughout your judicial career in criminal court?

PROBE: Do you manage misdemeanor cases? What types of misdemeanors? Approximately what percentage of your caseload do misdemeanors represent?

PROBE: Do you manage felony cases? What types of felonies? Approximately what percentage of your caseload do felonies represent?

5. Have you ever managed violent felony offenses?

IF YES: How would you describe the types of violent felony cases that you have handled?

PROBE: Have you handled homicide cases – approximately how many do you handle in an average year? Robbery cases? Assault cases? Rape cases?

6. Have you presided over trials as a judge in criminal court?

PROBE: Have you presided over bench trials?

IF YES: What types of cases? Approximately how many bench trials do you preside over in a given year?

PROBE: Have you presided over jury trials?

IF YES: What types of cases? Approximately how many jury trials do you preside over in a given year?

7. How would you describe the volume of cases you handle in a given year?

PROBE: Would you say that the volume of your caseload is below, equal or above average?

PROBE: Including pleas and trials, approximately how many cases do you manage in a given year?

8. How would you describe your legal experience prior to taking the bench?

PROBE: What kind of law did you practice? How long did you practice this type of law? Where did you practice?

PROBE: Did you practice criminal law as an attorney? Were you a prosecutor or a defense attorney? How long and where did you work in that capacity?

PERCEPTIONS OF JUDICIAL SENTENCING DISCRETION

Now I have some questions regarding your thoughts about judicial sentencing discretion.

9. What level of discretion do you believe is appropriate to vest with judges during the sentencing phase of a criminal prosecution?

PROBE: Do you ever feel as though you have too little discretion at sentencing?

IF YES: Under what circumstances?

PROBE: Do you ever feel that you have too much discretion at sentencing?

IF YES: Under what circumstances?

IF JUDGE HAS PRIOR EXPERIENCE AS AN ATTORNEY IN CRIMINAL CASES:

While practicing law as a criminal attorney, have you ever felt that judges have too much discretion in imposing sentence? Can you please describe the circumstances in which you felt that way?

Has your feeling changed now that you are a judge?

IF YES: In what way has your feeling changed?

10. What are your thoughts on sentencing guidelines?

11. Do you believe that sentencing guidelines should be instituted in all jurisdictions? Can you walk me through your thought process?

IF YES: Do you feel that guidelines should be compulsory or advisory? Can you walk me through the reason(s) that you feel this way?

PROBE: In your opinion, what are the challenges faced, if any, if sentencing guidelines are compulsory?

PROBE: In your opinion, what are the challenges faced if sentencing guidelines are advisory?

PROBE: What are your thoughts on mandatory minimum sentences?

PROBE: If you were asked to construct sentencing guidelines to help improve sentencing practices, how would you construct those guidelines?

SENTENCING FACTORS

Now I have some questions for you regarding the factors you consider when sentencing a defendant in a criminal case.

12. What factors do you consider when deciding the appropriate sentence to impose on a defendant in a criminal case?

13. Are there any factors that you think are more important than others?

IF YES: Which factors are the most important?

14. Are there certain factors that you consider to be irrelevant at sentencing?

IF YES: Which ones? Why do you think that these factors are irrelevant at sentencing?

15. What are your thoughts about aggravating circumstances as a sentencing consideration?

PROBE: Are there any aggravating factors that you deem to be more relevant than others? Which ones? Why do you feel that way?

16. What are your thoughts about mitigating circumstances as a sentencing consideration?

PROBE: Are there any mitigating factors that you deem to be more relevant than others? Which ones? Why?

17. Do you believe that judges should consider certain factors that are currently not identified as legally relevant?

IF YES: Can you describe those factors?

PERCEPTIONS OF EVIDENTIARY TYPE

Now I have some questions about your perceptions regarding different types of evidence.

18. Do you consider certain types of forensic evidence to be more reliable and probative than others?

IF YES: Which types?

PROBE: What are your thoughts regarding the reliability of forensic evidence produced in a state lab – such as fingerprint or ballistic evidence?

PROBE: What are your thoughts regarding the reliability of DNA evidence?

PROBE: What are your thoughts regarding the reliability of expert forensic evidence (e.g. identifying a defendant by a bite mark or footprint)?

PROBE: Have you presided over cases with state lab evidence?

IF YES: How often have you presided over cases with state lab evidence?

PROBE: Have you presided over cases with DNA evidence?

IF YES: How often have you presided over cases with DNA evidence?

PROBE: Have you presided over cases with expert forensic identification evidence (footprint, bite mark)?

IF YES: How often have you presided over cases with expert forensic identification evidence?

PROBE: In your opinion, how would you rank the reliability of state lab, DNA and expert forensic identification evidence?

PROBE: Is it beneficial for attorneys to use forensic evidence as part of their evidentiary package?

IF YES: What are the strengths of forensic evidence as an evidentiary form?

PROBE: Have you ever presided over a case where forensic evidence weakened the evidentiary package?

IF YES: Can you describe what happened in that case?

PROBE: Do you think that certain cases should have forensic evidence as part of the evidentiary package in order to produce a guilty verdict?

IF YES: What kinds of cases – in your opinion – should contain forensic evidence?

19. What are your thoughts regarding eyewitness testimony?

PROBE: What factors do you consider in determining whether a witness is credible or trustworthy?

PROBE: Which factors do you think are most important?

PROBE: Why do you think that these factors are most important in determining credibility?

PROBE: Have you ever concluded that a witness appearing in your courtroom lacked credibility?

IF YES: How often has this occurred?

PROBE: Can you provide me with an example or two of circumstances in which you concluded that a testifying witness lacked credibility?

PROBE: What factors do you consider in determining whether a witness is reliable or accurate?

PROBE: Which reliability factors do you think are most important?

PROBE: Why do you think that these reliability factors are most important?

PROBE: How often -- if ever -- have you felt that a witness appearing in your courtroom lacked reliability?

PROBE: Can you provide me with an example in which you concluded that a testifying witness lacked reliability?

PROBE: What are the strengths of eyewitness testimony as an evidentiary form?

PROBE: Have you ever presided over a case where an eyewitness weakened the evidentiary package?

IF YES: Can you describe what happened in that case?

PROBE: Do you think that certain cases should have eyewitnesses as part of the evidentiary package in order to produce a guilty verdict?

IF YES: What kinds of cases -- in your opinion -- should contain eyewitness evidence?

20. What are your thoughts regarding character witnesses?

PROBE: Do you think that character witnesses are an important piece of evidence? Can you walk me through your thought process?

21. What are your thoughts about alibi witnesses?

PROBE: Do you think that alibi witnesses are an important piece of evidence? Can you walk me through your thought process?

22. What are your thoughts about expert witnesses?

PROBE: Do you think that expert witnesses are an important piece of evidence? Can you walk me through your thought process?

23. In terms of probative value, how would you rank the following types of witnesses: eyewitness, character, alibi and expert? Can you walk me through your thought process?

24. What are the strengths and weaknesses of witness-based testimony as an evidentiary form?

25. Do you think that certain types/forms of evidence (i.e. forensic, witness-based) are stronger than others?

IF YES: Which types/forms? Can you explain your reasons?

PERCEPTIONS OF JURY VERDICTS

Now I have some questions about your perceptions of jury verdicts.

26. In general, how would you describe your thoughts about the jury system?

PROBE: Do you feel that juries are competent to judge the facts in a given case? Can you tell me a little more about why you feel this way?

PROBE: Based on your experience, do you feel that juries generally “get it right”? In what percentage of cases do they “get it right?”

PROBE: Do you think that juries generally consider the appropriate factors?

PROBE: What kinds of factors do juries consider most?

PROBE: Have you ever been dissatisfied by a jury verdict?

IF YES: How often have you been dissatisfied with a jury verdict?

PROBE: How would you describe the circumstances that result in this dissatisfaction?

PROBE: Do you have any recourse available in those situations?

IF YES: What recourse do you have?

27. Have you ever spoken to jury members after a trial was officially over?

IF YES: Did those conversations occur while you were a judge or an attorney? What was the purpose of those conversations?

PROBE: Without referencing any case specific facts, what types of questions did you ask?

PROBE: Did you feel satisfied with the answers that you received? Why (or why not)?

PROBE: Can you describe what, if anything, you learned during those conversations that you did not realize before?

Given the choice, would you prefer to preside over a jury trial or bench trial?

28. Can you tell me about the reason(s) for your preference?

29. Have you encountered any challenges while managing jury trials?

PROBE: Can you describe those challenges?

PROBE: How did you handle those challenges?

30. Do you believe that judges should have the power to set aside jury verdicts?

IF YES: Under what circumstances do you believe this should be the case?

PROBE: Have you ever set aside a jury's guilty verdict?

IF YES: Can you provide me with a few examples?

31. Have you ever felt uncertain of your own verdict at a bench trial? What were the reasons for your uncertainty? How did you handle that uncertainty?

INDIVIDUALIZED SENTENCING

Now I have some questions about your views on individualized sentencing.

32. Do you think that sentencing differences between offenders convicted of similar offenses and possessing similar criminal histories are a source of concern? Can you explain your viewpoint?

33. Do you think that sentences should be customized to individualized defendants?

IF YES: To what extent do you think that sentences should be customized?
How much variation is appropriate?

34. In determining the appropriate sentence to impose upon a convicted defendant, how important is protection of the community as a sentencing consideration? Why do you feel this way?

35. How important is the blameworthiness of the defendant as a sentencing consideration? Why do you feel this way?

36. Do practical concerns – such as whether the defendant is responsible for minor children – influence the type or length of sentence that you impose in a given case?

IF YES: To what extent do practical concerns influence the type or length of sentence that you impose in a given case?

VIGNETTES

As I explained to you when we set up this interview, I am studying the factors that judges consider in deciding the sentence to impose on a defendant in a criminal case. With your permission, I would like to present you with a written description of 4 hypothetical case scenarios. Each hypothetical case describes a criminal act and provides you with certain defendant and other information. I would like you to take a few moments to read each hypothetical and then write down, based on the information provided, the sentence that you would

impose in each case. For purposes of this exercise, please assume that the information provided represents the full universe of information available in the case.

- 37. Can you walk me through the process by which you reached your decision in Case A?
- 38. Can you walk me through the process by which you reached your decision in Case B?
- 39. Can you walk me through the process by which you reached your decision in Case C?
- 40. Can you walk me through the process by which you reached your decision in Case D?
- 41. I see that you sentenced the defendant in Case ____ to the longest sentence and the defendant in Case ____ to the shortest sentence. Can you tell me about the reasons for the differences in the two sentences?

EVIDENTIARY WEIGHT/CONFIDENCE IN GUILT

- 42. Do you consider the strength of the evidence in determining sentence?
- 43. In what way do your own perceptions of guilt and/or the accuracy of the jury's verdict factor into your sentencing decisions?

WRAPPING UP

We are at the end of our interview. Before we conclude, is there any other information that you would like to share?

As I told you, the purpose of this study is to better understand the factors that judges consider during the sentencing process. Are there any areas, not covered by this interview, that you feel would be helpful in future interviews.

Thank you so much for your time.

Appendix 3.3: Vignettes

Scenario A

The defendant was convicted of Assault in the First Degree (in New York) after a jury trial.

The complainant (victim) stated that he was at a bar having a beer when he was approached by the defendant. The complainant did not know the defendant prior to this incident. According to the complainant, the defendant ordered him to move to a different bar stool. The complainant refused. The defendant then pulled out a knife and stabbed the complainant in the leg. After the incident, the complainant told the bartender (who had stepped away during the incident) what happened and the bartender called the police and an ambulance. When the police arrived 10 minutes later, the complainant pointed out the defendant and stated that the defendant was the person who stabbed him. The complainant later received twelve stitches in the hospital. There were approximately ten other people in the bar at the time that this incident occurred. However, upon questioning by the police, all ten individuals claimed that they did not see the incident. Additionally, the knife was not recovered.

The defendant has denied all allegations.

The jury found the complainant credible and convicted defendant.

This is the defendant's first arrest.

Scenario B

The complainant was convicted of Assault in the First Degree (in New York) after a jury trial.

The complainant stated that he was at a party, just hanging out, when the defendant approached him. The defendant accused the complainant of making advances toward the defendant's girlfriend. The complainant denied making such advances and told the defendant that he was paranoid. The defendant and the complainant did not know each other prior to this incident. The defendant then punched the complainant in the stomach, removed a small knife from his pocket and stabbed the complainant in the arm. The complainant needed 8 stitches to close his wound but otherwise sustained only minor bruises. There were three eyewitnesses that came forward and corroborated the complainant's version of events. Additionally, a knife was recovered that tested positive for blood. The blood type on the knife matched the complainant's blood type.

The defendant did not make any statements.

This is the defendant's first arrest.

Scenario C

Defendant was convicted (in New York) of Robbery in the First Degree.

According to the complainant, she was on her way home from a friend's house when the defendant approached her. The defendant asked her whether he could borrow her cell phone to make a phone call. It was daylight. The complainant handed over her cell phone and the defendant told her that he was going to keep it. The complainant tried grabbing it back and the defendant, who had been dismantling boxes at his job, pulled out a box cutter and slashed the complainant's hand. The defendant then ran away. The complainant flagged down a police car that was in the vicinity. The complainant identified the defendant ten blocks away, 5 minutes later. A man -- who had been standing across the street at the time that the incident occurred -- came forward and stated that he saw what happened and was prepared to cooperate with the police. He came down to the precinct and observed a lineup. He identified the defendant in the lineup. The box cutter was not recovered and the defendant did not make any statements. This incident is the defendant's first arrest.

Scenario D

The defendant was convicted of Robbery in the First Degree (in New York) after a jury trial.

According to the victim, the defendant approached her as she was walking home from work. No one else was on the street at that time. It was still daylight. The defendant pulled out a knife, pointed it at the victim and demanded money. The victim handed over \$60. The defendant then grabbed a gold chain from the victim's neck, breaking the clasp. The defendant grabbed the money, slashed the victim on her arm, dropped the knife and fled from the location. The victim and the defendant did not know each other prior to this incident. The entire encounter lasted less than 2 minutes. There were no other witnesses.

The victim dialed 911 on her cell phone. The police responded to the scene 10 minutes later. The victim – who was upset but composed – gave a description of the assailant. The victim also showed police the knife that was on the ground that had been used by her assailant. The police vouchered the knife. The police placed the victim in the police car and drove around the area. Approximately 8 blocks away (20 minutes after the incident), the victim observed the defendant walking on the street. She identified him as the man who had robbed her. Subsequent testing suggested that the knife contained fingerprints that belonged to the defendant. The gold chain was not recovered. The victim was treated for her injury with stitches and antibiotics. The defendant refused to make any statements to the police and did not testify at trial.

This is the defendant's first arrest.