3-2016

Popularizing Hearsay

Justin Sevier
Florida State University College of Law

Popularizing Hearsay

JUSTIN SEVIER*

TABLE OF CONTENTS

* Assistant Professor, Florida State University College of Law. Ph.D. Candidate in Psychology, Yale University (2015); J.D., Harvard Law School (2006). I thank Tom R. Tyler, Dan Kahan, Kip Viscusi, Joni Hersch, Arden Rowell, Kenworthey Bilz, Adam Kolber, Christina Mulligan, Brian Lee, Jack Chin, Alex Stein, Michael Saks, Mark Spottswood, Mark Seidenfeld, the Yale University Department of Psychology, The Florida State University College of Law faculty, and the editors of *The Georgetown Law Journal* for their insightful oral and written comments with respect to the topics discussed in this Article. I also thank Vincent Fernandez and Amanda Qadri for excellent research assistance.

INTRODUCTION

> When first the Queene was crown'd, I lived well;
> Butt Spaniards came, and Wyatt quicklie rose;
> Whoe onelie meant outlandish force to quell:
> To him adherent men didd mee suppose.
> When hee had play'd his woefull tragedie;
> Then next ensued my bitter comedie.[1]
>
> —The Ghost of Nicholas Throckmorton (1874)

In the spring of 1554, the English Crown charged Nicholas Throckmorton, a member of the Warwickshire gentry, with high treason for his alleged role in the Wyatt Rebellion.[2] Several months earlier Queen Mary Tudor, a Roman Catholic, had risen to power and intended to marry Prince Phillip II of Spain.[3]

---

1. *Throckmorton's Ghost*, THE LEGEND OF SIR NICHOLAS THROCKMORTON 29 (John Gough Nichols ed., 1874).

2. *See* D. M. LOADES, TWO TUDOR CONSPIRACIES 95–97 (1965); *see also* THE TRIAL OF NICHOLAS THROCKMORTON (Annabel Patterson ed., 1998) (reprinting the trial in full and including annotations); 4 RAPHAEL HOLINSHED, *The Trial of Sir Nicholas Throckmorton*, *in* THE CHRONICLES OF ENGLAND, SCOTLAND, AND IRELAND 1737 (London, 1577) (reprinting the trial record without annotations). According to the Crown, a faction of English noblemen planned to stage uprisings in several English counties before converging on London to replace Queen Mary with her half-sister Elizabeth. *See* LOADES, *supra* at 15–19. One historian has theorized that the true purpose of Throckmorton's treason trial was to implicate Elizabeth, among others. *See id*. at 96.

3. *See* DAVID LOADES, THE REIGN OF MARY TUDOR: POLITICS, GOVERNMENT, AND RELIGION IN ENGLAND 1553–1558, at 57–95 (2d ed., 1991) (discussing the turbulent political climate that existed at the outset of Mary Tudor's rule, and detailing the events leading to her marriage to Prince Phillip II of Spain). Mary Tudor eventually married Phillip II in a lavish ceremony. *See id*. at 157–58 (noting that "[t]heir

Powerful Protestant forces among the English highborn sought to ensure that Queen Mary's reign—and her marriage to Prince Phillip—was short-lived.[4] Soon after the Queen's ascension to the throne, Sir Thomas Wyatt of Kent and several coconspirators allegedly hatched a plot, depending upon whom one believes, to implore the Queen to marry an Englishman instead, or to assassinate her.[5]

If convicted of high treason for his role in the rebellion, Throckmorton faced execution at the Tower of London.[6] As he faced his triers at Guildhall, he pleaded to the tribunal to bring to him, "face to face to justify the matter," with the man whose confession brought Throckmorton to trial.[7]

Throckmorton's trial at Guildhall was notable in one fundamental respect: the Crown's case against him for high treason consisted almost entirely of hearsay evidence.[8] The Crown presented as the centerpiece of its case against Throckmorton the deposition of the Duke of Suffolk, who did not testify at trial.[9] This deposition included several hearsay statements that the Duke of Suffolk allegedly heard from his brother,[10] which suggested that Throckmorton "was privy to

---

wedding, which was celebrated on St James day (25 July) in Winchester cathedral, was a ceremony of high emotional and political excitement").

4. *See* JOHN BELLAMY, THE TUDOR LAW OF TREASON: AN INTRODUCTION 55 (1979) ("According to the indictment, Throckmorton, with ten other gentlemen and in the company of certain traitors, had compassed to deprive the queen of her crown and dignity and destroy her, and to take the Tower of London and levy war against her."); *see also* THE TRIAL OF NICHOLAS THROCKMORTON, *supra* note 2, at 7 (discussing how "Protestant courtiers . . . tried desperately and unsuccessfully to prevent [King Edward VI's] Roman Catholic elder sister, Mary Tudor, from succeeding him," and asserting that the Wyatt Rebellion was directed "against the new queen, and particularly against her planned marriage to Philip II of Spain.").

5. *See* BELLAMY, *supra* note 4, at 55 (detailing the plot "to deprive the queen of her crown and dignity and destroy her" and "[c]ompassing to deprive the prince of his crown"). *See generally* THOMAS DEKKER & JOHN WEBSTER, THE FAMOUS HISTORY OF SIR THOMAS WYAT (1607) (providing a detailed chronology of Sir Thomas Wyatt and discussing the circumstances surrounding Wyatt's Rebellion).

6. *See* BELLAMY, *supra* note 4, at 182 (explaining that "general governmental policy [in Tudor-era England] was either to execute convicted traitors forthwith, or release them fairly soon after arraignment").

7. *See* H.B. WILSON, THE AMERICAN JUROR: BEING A GUIDE FOR JURYMEN THROUGHOUT THE UNITED STATES 192–95 (Philadelphia, J.B. Lippincott & Co. 1868) (highlighting important events from the trial and recounting the exchange between Throckmorton and the Crown). Guildhall was the venue in which ceremonies, trials, and other public events took place in London during the Tudor era. It still stands today. *See Guildhall*, CITY OF LONDON, http://www.guildhall.cityoflondon.gov.uk/ (last visited Nov. 15, 2015).

8. *See* JAMES BRADLEY THAYER, A SELECTION OF CASES ON EVIDENCE AT THE COMMON LAW 313 (2d ed. 1900) (quoting from Sir Walter Raleigh's trial for treason, in which Raleigh expresses concern that he "may be massacred by mere hearsay, as Sir Nicholas Throckmorton was like to have been in Queen Mary's time").

9. *See* 1 JAMES FITZJAMES STEPHEN, A HISTORY OF THE CRIMINAL LAW OF ENGLAND 328 (London, MacMillan & Co. 1883) ("A deposition of the Duke of Suffolk was next read, on which Throckmorton remarked that it stated only what the Duke said he had heard from his brother . . . .").

10. *Id.*

the whole devices against the Spaniards."[11]

By all historical accounts, the tribunal was openly hostile to Throckmorton,[12] and he was not allowed to cross-examine the Duke of Suffolk nor Suffolk's brother.[13] Nonetheless, the defendant eloquently urged the jury not to convict him on evidence from accusers who did not look him in the eye at trial.[14] Although many of the alleged conspirators involved in the Wyatt Rebellion were convicted,[15] Throckmorton's appeal to the jury was successful, and the jury acquitted him.[16] The verdict so outraged the Crown that it levied heavy fines against the jurors and imprisoned them.[17]

The historical record from this trial does not answer one important question: why did the jury decide to acquit Throckmorton when other English juries convicted nearly all of his contemporaries? Although the historical record suggests that the jury refused to convict Throckmorton because of the unfairness of the procedures used by the Crown,[18] including its reliance on hearsay evidence,[19] the basis for the jury's dissatisfaction with hearsay remains a mystery.

The hearsay rule, a complex doctrine that purports to bar secondhand evidence in court, has been characterized as a "spoiled child" by academic scholars because of the disproportionate attention that it has received compared to other principles of evidence.[20] Despite the wealth of research that attempts to add clarity to this increasingly complex doctrine, legal scholars and policymakers

---

11. 1 COBBETT'S COMPLETE COLLECTION OF STATE TRIALS AND PROCEEDINGS FOR HIGH TREASON AND OTHER CRIMES AND MISDEMEANORS FROM THE EARLIEST PERIOD TO THE PRESENT TIME 883 (London, R. Bagshaw et al., 1809).

12. *See, e.g.*, THE TRIAL OF NICHOLAS THROCKMORTON, *supra* note 2, at 8 (explaining that Throckmorton's trial was one of several "exemplary political trials, of considerable importance" to the Crown, and noting that the jurors that acquitted Throckmorton "were subsequently imprisoned in the Tower and the Fleet, and fined exorbitantly" for their verdict).

13. *See* STEPHEN, *supra* note 9, at 328 (reprinting the portion of the trial where the Duke of Suffolk's hearsay testimony was presented to the jurors in the Duke of Suffolk's absence).

14. LOADES, *supra* note 2, at 97 (classifying Throckmorton's performance at his trial as "a very able defence," noting its ten-hour duration, and stating that Throckmorton's acquittal led to "rejoicings in the City"); *see also* THE TRIAL OF NICHOLAS THROCKMORTON, *supra* note 2, at 20 (noting that "the drama of the queen vs. Sir Nicholas Throckmorton unfolded . . . as a potential tragedy transformed by the wit and intelligence of the protagonist into a political comedy of manners").

15. *See* THE TRIAL OF NICHOLAS THROCKMORTON, *supra* note 2, at 7 (noting that "Wyatt and other military leaders of the conspiracy were, of course, tried and executed" along with "Mary's Protestant rival, Lady Jane Grey"). *But see* LOADES, *supra* note 2, 110–14 (noting that other members of the various uprisings escaped execution).

16. *See* LOADES, *supra* note 2, at 97. Many other famous English defendants were not so fortunate. *See* THAYER, *supra* note 8, at 313 (discussing the trials of Sir Walter Raleigh, William Thomas, and Sir Matthew Hale).

17. *See* LOADES, *supra* note 2, at 97.

18. *See* THE TRIAL OF NICHOLAS THROCKMORTON, *supra* note 2, at 19–20 (noting that modern readers would "be puzzled—not to say shocked—" by "striking feature[s]" of the trial, which included denying Throckmorton legal counsel and other rights guaranteed by the United States Constitution).

19. *See* THAYER, *supra* note 8, at 313 (including a quote from Sir Walter Raleigh at his trial, suggesting Raleigh's belief that Throckmorton was tried primarily on hearsay evidence).

20. JOHN H. WIGMORE, A STUDENTS' TEXTBOOK OF THE LAW OF EVIDENCE 238 (1935); Stephan Landsman & Richard F. Rakos, *Research Essay: A Preliminary Empirical Enquiry Concerning the*

are at odds with respect to its rationale. Although the rationales put forth vary, they fall into two camps. The primary rationale put forth by scholars and rule makers is that the rule prohibiting hearsay evidence promotes decisional accuracy by barring unreliable evidence from the courtroom.[21] Others say, however, that the hearsay rule promotes norms of procedural justice by disallowing evidence from accusers who will not, as the expression goes, "look the accused in the eye."[22]

The primary rationale for the hearsay rule suffers from two important limitations. First, if the removal from trials of potentially unreliable evidence will decrease the risk of mistaken verdicts, it must be true that judicial fact finders will place more emphasis on hearsay than it deserves, relative to its actual probative value. Empirical research on how fact finders evaluate hearsay evidence, however, casts serious doubt on this premise.[23] Second, although the empirical research to date casts doubt on the premise, it is actually very difficult to measure how laypeople respond to hearsay, which makes definitive statements regarding the role of the hearsay rule on the accuracy of verdicts difficult to support.

The minority rationale—that the hearsay rule promotes litigants' dignity interests and increases the legitimacy of the tribunal to the public—has the support of social psychological research and would add coherence to the hearsay doctrine, but it currently suffers from a lack of empirical data. No one has yet examined why laypeople appear to dislike hearsay evidence and why they believe the rule against hearsay is in place.

This Article is the first to examine how laypeople respond to hearsay evidence and is the first to test these competing theories against one another. This Article hypothesizes that, like the jury who decided the fate of Nicholas Throckmorton, laypeople will be dissatisfied with trials that include hearsay evidence compared to trials that involve live witnesses who are subject to

---

Prohibition of Hearsay Evidence in American Courts, 15 LAW & PSYCHOL. REV. 65, 65 (1991) (mentioning this trend).

21. The first scholarly articulation of the decisional accuracy rationale was made by Professor Edmund Morgan of Harvard Law School in the early 20th century. *See* Edmund M. Morgan, *Some Suggestions for Defining and Classifying Hearsay*, 86 U. PA. L. REV. 258, 258–59 (1938) (noting that, in crediting an in-court witness, the trier of fact evaluates the witness's "perception, recollection, narration and veracity" and explaining that crediting an in-court witness's recounting of a hearsay statement poses greater challenges to the fact finder). This rationale has been articulated in common law opinions, however, since at least the 1800s. *See* Roger Park, *The Hearsay Rule and the Stability of Verdicts: A Response to Professor Nesson*, 70 MINN. L. REV. 1057, 1058 nn.5–6 (1986) (collecting cases from England and the United States, dating back as far as 1812, and discussing the dangers of hearsay evidence in terms of its effects on a fact finder's accuracy).

22. The scholarly articulation of a fairness rationale for the rule barring hearsay evidence first appeared explicitly in the 1980s. *See* Roger Park, *A Subject Matter Approach to Hearsay Reform*, 86 MICH. L. REV. 51 (1987); Eleanor Swift, *A Foundation Fact Approach to Hearsay*, 75 CALIF. L. REV. 1339 (1987). See *infra* Part II for a more robust discussion of procedural justice.

23. *See infra* note 54 and accompanying text. *See also* Justin Sevier, *Testing Tribe's Triangle: Juries, Hearsay, and Psychological Distance*, 103 GEO. L.J. 879, 893–96 (2015) (discussing over a dozen studies that challenge the proposition that juries overvalue hearsay evidence).

cross-examination. Moreover, this Article hypothesizes, based on research on a psychological phenomenon called procedural justice, that participants prefer the hearsay rule because the rule affords defendants dignity and respect by requiring their accusers to look them in the eye.[24] The Article will then argue that a procedural justice rationale for barring hearsay evidence will lead to a more coherent hearsay rule that enjoys the legitimacy of the public at large.

This Article serves as a sequel to my previous work, *Testing Tribe's Triangle: Juries, Hearsay, and Psychological Distance*.[25] In that previous work, I tested whether the prevailing rationale for the rule barring hearsay evidence—that jurors are likely to overvalue hearsay because the hearsay declarant is often unavailable for cross-examination—withstands empirical scrutiny. This Article builds on the conclusions reached in that previous work and explores a more empirically sound rationale for the hearsay rule: that it increases popular perceptions of the procedural justice afforded by the courts because it purports to bar second-hand evidence from the courtroom.

Part I of this Article (1) discusses the conclusions reached in *Testing Tribe's Triangle: Juries, Hearsay, and Psychological Distance* and (2) argues that the prevailing rationale for the rule barring hearsay evidence—that it promotes decisional accuracy—is both empirically suspect and difficult to measure meaningfully. Part II introduces an alternative rationale for the hearsay rule: procedural justice. It also argues that a procedural justice rationale for the rule barring hearsay evidence is superior to the prevailing rationale for the hearsay rule, and that it is more likely to cultivate popular legitimacy for the rule. Parts III and IV report the results from two original experimental studies that examine laypeople's understandings about, and attitudes toward, hearsay evidence. Part V explores the implications of these findings to hearsay policy, responds to potential objections raised regarding the methodology of these studies and their implications, and concludes that rule makers and scholars should fashion the hearsay rule so that it explicitly promotes the interests of procedural justice and fair play.

I. *TESTING TRIBE'S TRIANGLE*: THE INFERIORITY OF DECISIONAL ACCURACY

Because this Article builds upon—and serves as a sequel to—my prior work in *Testing Tribe's Triangle: Juries, Hearsay, and Psychological Distance*,[26] this

24. Procedural justice is defined as the perceived fairness of the process that is used to resolve disputes and allocate resources. *See* Lawrence B. Solum, *Procedural Justice*, 78 S. CAL. L. REV. 181, 183 (2004). *See generally* E. ALLAN LIND & TOM R. TYLER, THE SOCIAL PSYCHOLOGY OF PROCEDURAL JUSTICE (1988) (reviewing theories on procedural justice and implications for legal, political, interpersonal, and work-related settings); JOHN THIBAUT & LAURENS WALKER, PROCEDURAL JUSTICE: A PSYCHOLOGICAL ANALYSIS (1975) (comparing whether adversary or inquisitorial procedures better serve procedural justice); TOM R. TYLER, WHY PEOPLE OBEY THE LAW (2006) (comparing the instrumental and normative perspectives of procedural justice). For an in-depth discussion of the role of procedural justice in the legal system, see *infra* Part II.

25. Sevier, *supra* note 23.

26. *Id.*

Part will discuss the central argument raised in that work, the empirical test that was conducted, and the conclusions that were reached with respect to the prevailing rationale for the rule barring hearsay evidence. This Part will also discuss the research questions that remained unanswered in my prior work and how this Article addresses those questions.

A. EMPIRICAL WEAKNESSES

Testing Tribe's Triangle: Juries, Hearsay, and Psychological Distance examined the prevailing rationale for the rule barring hearsay, which refers to out-of-court (often second-hand) statements that are proffered as evidence for the truth of the matters asserted in the statements.[27] My prior work noted that, although courts purport to ban hearsay evidence when it is used for that purpose, courts recognize—depending on how one counts—thirty-one different exceptions to this rule, in which statements that, conceptually, are hearsay are admitted into evidence for the truth of the matters asserted therein.[28]

Although policymakers and scholars have hypothesized several different rationales for the hearsay rule—in a manner reminiscent of a Whac-a-Mole carnival game[29]—they break down into two camps: a prevailing rationale and a minority rationale.[30] The prevailing rationale for the rule barring hearsay evi-

---

27. *See* FED. R. EVID. 801(c) ("'Hearsay' means a statement that: (1) the declarant does not make while testifying at the current trial or hearing; and (2) a party offers in evidence to prove the truth of the matter asserted in the statement."); FED. R. EVID. 801(a) ("'Statement' means a person's oral assertion, written assertion, or nonverbal conduct, if the person intended it as an assertion."); FED. R. EVID. 802 ("Hearsay is not admissible unless any of the following provides otherwise: a federal statute; these rules; or other rules prescribed by the Supreme Court.").

28. *See* FED R. EVID. 801(d) (allowing into evidence hearsay statements from opposing parties in litigation, statements of coconspirators, and prior statements from testifying witnesses); FED R. EVID. 803 (allowing into evidence hearsay statements that qualify as present sense impressions; excited utterances; statements made for the purpose of seeking medical diagnosis or treatment; business records; public records; ancient documents; statements of then-existing mental, emotional, or physical condition; family records; recorded recollections; and learned treatises, among others); FED. R. EVID. 804 (allowing into evidence, when the declarant in unavailable, hearsay statements that qualify as dying declarations; statements of personal or family history; former testimony; statements against interest; and statements from declarants who became unavailable through a party's wrongful actions); FED. R. EVID. 807 (allowing into evidence hearsay statements that do not qualify for admissibility under FED. R. EVID. 801, 803 or 804 if, among other criteria, the statement "has equivalent circumstantial guarantees of trustworthiness").

The United States Supreme Court has, however, limited the admissibility of hearsay statements used against a defendant in a criminal trial if those otherwise admissible statements violate the Confrontation Clause of Article VI of the United States Constitution. *See* Crawford v. Washington, 541 U.S. 36, 68–69 (2004) (stating that "the only indicium of reliability sufficient to satisfy constitutional demands is the one the Constitution actually prescribes: confrontation"); *see also* Michigan v. Bryant, 562 U.S. 344, 348–49 (2011) (clarifying the types of statements that are testimonial and, therefore, invoke the Confrontation Clause); Giles v. California, 554 U.S. 353, 358 (2008) (same); Davis v. Washington, 547 U.S. 813, 822 (2006) (same).

29. *See* Sevier, *supra* note 23, at 890 ("The rationale for the rule barring hearsay is a moving target at best and, at worst, a futile game of policy Whac-A-Mole.").

30. This Article does not argue that the prevailing and minority rationales that scholars have proffered to justify the existence of the rule barring hearsay evidence are mutually exclusive or

dence argues that second-hand hearsay evidence should be excluded because it lacks the reliability of in-court testimony, which is tested under oath, through cross-examination, in an environment in which jurors observe the witness's demeanor while testifying.[31] Because out-of-court statements are often not presented to a jury under those circumstances, evidence policymakers believe that there is a substantial risk that fact finders will overvalue those out-of-court statements.[32]

In his influential article, *Triangulating Hearsay*, Professor Laurence Tribe of Harvard Law School succinctly conceptualized the prevailing rationale for the hearsay rule, which focuses primarily on the efficacy of cross-examination.[33] Professor Tribe noted that common law judges recognized four capacities that affect all testimony presented by witnesses in court: the witness's narrative clarity, sincerity, perception of the event to which the witness testifies, and the witness's memory for the event.[34] Professor Tribe then conceptualized these capacities in terms of a "testimonial triangle."[35] His triangle, illustrated below in my prior work, raises two questions that fact finders must ask of witnesses: (1) Does the witness's in-court statement reflect her true belief about the event? and (2) if so, does the witness's true belief about the event reflect what actually occurred?[36] Professor Tribe noted that the first question—the left leg of the

---

incompatible. Indeed, the U.S. Supreme Court majority's holding in *Crawford v. Washington* implicates more than one single rationale for the rule barring hearsay evidence. *See Crawford*, 541 U.S. at 50–54. This Article examines which of these rationales should be the primary rationale for the doctrine, which in turn, will have implications for the future direction of the hearsay rule.

31. *See, e.g.*, JAMES F. BAILEY, III & OSCAR M. TRELLES, II, *Introductory Note: The Hearsay Problem*, *in* 2 THE FEDERAL RULES OF EVIDENCE LEGISLATIVE HISTORIES AND RELATED DOCUMENTS 152, 153 (1980) (noting that "[t]he belief, or perhaps hope, that cross-examination is effective in exposing imperfections of perception, memory, and narration" underlies the rule barring hearsay evidence, and that some hearsay is admitted into evidence under the current hearsay scheme when the circumstances surrounding the evidence "furnish guarantees of trustworthiness"); *see also* Edmund M. Morgan, *Hearsay Dangers and the Application of the Hearsay Concept*, 62 HARV. L. REV. 177 (1948) (espousing the common law objections to hearsay evidence); Frederick Schauer, *On the Supposed Jury-Dependence of Evidence Law*, 155 U. PA. L. REV. 165, 167–68 (2006); Laurence H. Tribe, *Triangulating Hearsay*, 87 HARV. L. REV. 957 (1974) (describing the common law objections to hearsay and conceptualizing them into a "[t]estimonial [t]riangle"). This rationale is also explicitly stated by evidence rule makers in the Advisory Committee's Note accompanying Federal Rule of Evidence 803. The Advisory Committee's Note argues, without citing empirical evidence, that "[t]he present rule proceeds upon the theory that under appropriate circumstances a hearsay statement may possess circumstantial guarantees of trustworthiness sufficient to justify nonproduction of the declarant in person at the trial even though he may be available." FED. R. EVID. 803 advisory committee's note; *see also* Park, *supra* note 21, at 1057 ("The conventional explanation for the exclusion of hearsay centers on the danger of admitting evidence that has not been tested for reliability. Unlike courtroom witnesses, hearsay declarants have not testified under oath, in the presence of the trier of fact, and subject to cross-examination.").

32. BAILEY & TRELLES, *supra* note 31, at 153.

33. Tribe, *supra* note 31.

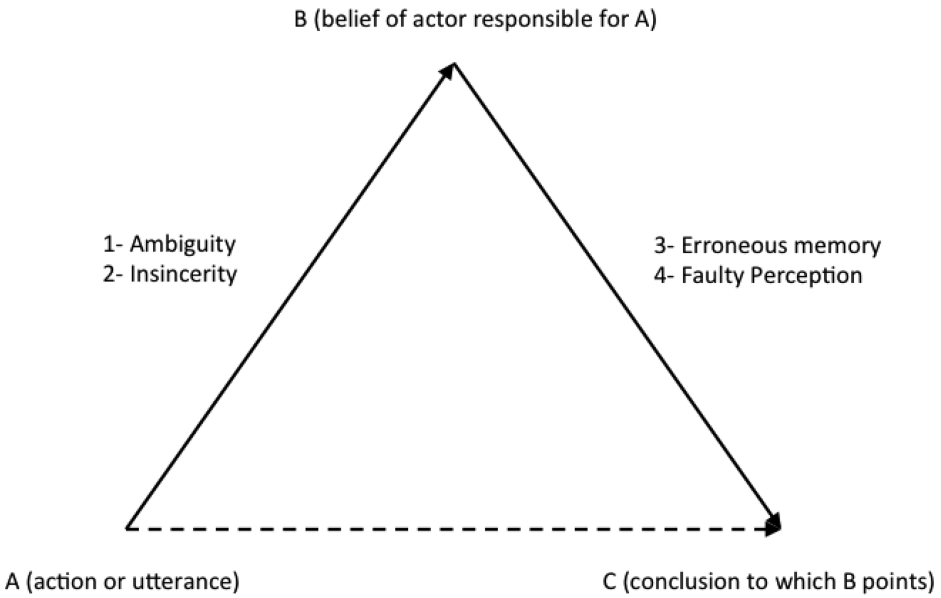
34. *See id*. at 958. These testimonial capacities are now embodied in the personal knowledge requirement of Federal Rule of Evidence 602 and the oath or affirmation that witnesses undertake pursuant to Federal Rule of Evidence 603. *See* DEBORAH JONES MERRITT & RIC SIMMONS, LEARNING EVIDENCE: FROM THE FEDERAL RULES TO THE COURTROOM 160 (3d ed. 2014).

35. Tribe, *supra* note 31, at 959.

36. *See id*.

triangle—depends on the witness's narrative clarity and sincerity.[37] The second question—the right leg of the triangle—depends on the witness's accurate perception of, and memory for, the event to which she testifies.[38] A failure of any of these testimonial capacities—which should be exposed upon cross-examination of the witness—should cause the fact finder to decrease the probative value that she assigns the witness's testimony.[39]

Figure 1: The Testimonial Triangle[40]



The difficulty with hearsay evidence, according to the reliability rationale for excluding it, is that hearsay testimony is often subject to two testimonial triangles: the testimonial triangle of the in-court witness and the triangle belonging to the out-of-court hearsay declarant. And, most importantly, the testimonial triangle that belongs to the out-of-court hearsay declarant is difficult, if not impossible, to meaningfully evaluate through cross-examination.[41] Thus, the reliability rationale states that hearsay evidence contains the risk of overvaluation by the fact finder, who is likely to place too much probative

37. *See id*.
38. *See id*.
39. *See id*. at 958–61.
40. This figure originally appeared in my article *Testing Tribe's Triangle: Juries, Hearsay, and Psychological Distance*. Sevier, *supra* note 23, at 893.
41. The Federal Rules of Evidence do allow parties, to the extent possible, to impeach the credibility of out-of-court declarants. *See* FED. R. EVID. 806 (stating that, when a hearsay statement "has been admitted in evidence, the declarant's credibility may be attacked, and then supported, by any evidence that would be admissible for those purposes if the declarant had testified as a witness").

weight on risky, potentially unreliable evidence.[42]

This is, of course, an empirical claim about how fact finders—and in particular, jurors—reach decisions regarding how to weigh evidence. And although social psychological evidence strongly suggests that jurors are skeptical of hearsay evidence,[43] the premise of the testimonial triangle—insofar as it embodies the prevailing rationale for the rule barring hearsay evidence—had not yet been tested directly.

In *Testing Tribe's Triangle: Juries, Hearsay, and Psychological Distance*, I tested this rationale empirically through the lens of the social psychological phenomena termed "construal level theory."[44] In two studies reported in this prior work, I presented laypeople with trial vignettes in which I manipulated (1) the type of testimonial capacity at issue in the case; and (2) the identity of the individual whose capacity was at issue—either the in-court witness's capacity, the out-of-court hearsay declarant's capacity, neither, or both.[45] I then examined the probative weight that the participants had placed on the evidence.[46]

The data that I collected challenged the view—implicit in the prevailing rationale for the hearsay rule—that jurors systematically overvalue hearsay evidence. Instead, I found that jurors tend to be skeptical of hearsay evidence, and they do not give it the same weight that they give in-court testimony that is subject to cross-examination.[47] Further, I found that jurors were sensitive to weaknesses in an individual's testimonial capacities, and not just with respect to in-court witnesses, but with respect to out-of-court hearsay declarants as well.[48]

In a follow-up experiment, I examined how mock jurors respond to "hearsay within hearsay," which occurs when a hearsay statement itself also contains hearsay.[49] If jurors discount out-of-court statements because such statements are not subject to the scrutiny of cross-examination, then we would expect hearsay statements that contain additional levels of hearsay to be even less persuasive to jurors than "ordinary" hearsay statements.[50] And, indeed, the data suggested

42. *See supra* note 21.

43. *See* Sevier, *supra* note 23, at 893–96 (discussing, in detail, the results from over a dozen empirical studies examining mock jurors' attitudes toward hearsay and their decisions in cases that involve hearsay). The experiments strongly suggest that jurors discount hearsay evidence sharply and that they pay attention to cognitive and motivational factors that threaten to decrease the probative value of hearsay. *See id.*

44. *See* Sevier, *supra* note 23, at 896–900 (explaining the tenets of construal level theory and discussing several experiments in which construal level theory has been tested).

45. *Id.* at 903–15.

46. *Id.*

47. *Id.* at 914–15 (discussing the results of Study 1).

48. *Id.*

49. *Id.* at 915–22. Hearsay statements that are embedded within other hearsay statements are the subject of Federal Rule of Evidence 805: "Hearsay within hearsay is not excluded by the rule against hearsay if each part of the combined statements conforms with an exception to the rule." FED. R. EVID. 805.

50. *See* Sevier, *supra* note 23, at 916 (discussing the hypothesis that "jurors are most persuaded by non-hearsay testimony, less persuaded by hearsay, even less persuaded by double hearsay, and least persuaded by triple hearsay testimony").

that jurors systematically devalue out-of-court statements consistent with the degree of hearsay that is contained in those statements.[51] In sum, the studies suggest that jurors are more careful in evaluating hearsay statements than evidence policymakers believe, and the data suggest that the prevailing rationale for the rule barring hearsay evidence—that jurors will overvalue risky, potentially unreliable evidence—may not be empirically valid.[52]

B. THE MEASUREMENT PROBLEM

Although it is discussed in passing in my prior work, this Article expands on another difficulty with respect to the prevailing rationale for the hearsay rule. To the extent that scholars can measure the extent to which jurors weigh hearsay evidence, it appears that jurors weigh it more carefully than evidence policymakers have acknowledged.[53] But a second argument that challenges the reliability rationale for the hearsay rule stems from an entirely different perspective: because it is difficult—or even impossible—to truly measure whether jurors give appropriate weight to hearsay evidence, the decisional accuracy rationale for the hearsay rule is an unfalsifiable empirical prediction regarding how jurors respond to hearsay.

Empirical research contributes significantly to our understanding of how laypeople process and weigh hearsay information. These studies suggest, strongly, that jurors do not afford hearsay evidence the same weight that they afford in-court testimony.[54] But it is far more difficult to argue that these studies show that laypeople act rationally with respect to how they process hearsay evidence, or that laypeople afford hearsay evidence its appropriate weight.

To determine that jurors make rational decisions with respect to hearsay evidence, we require data showing not only that jurors discount hearsay evidence compared to nonhearsay, but that they also discount that evidence by the correct amount. This normative claim is difficult to argue because it presupposes that the meaning—and probativeness—of a piece of evidence has a fixed value that can be measured reliably.[55] Unfortunately, assessing the probative value of evidence is a topic that has vexed legal scholars for decades; there is

---

51. *See id.* (discussing the results of Study 2).

52. *See id*. at 922–28 (discussing the results from Study 1 and Study 2 and exploring their implications).

53. *See supra* notes 47–51 and accompanying text.

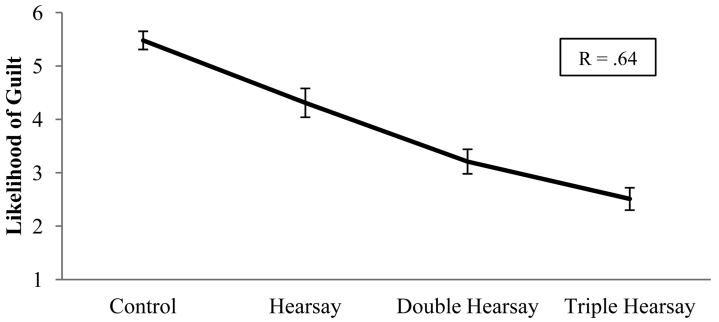
54. *See, e.g*., Landsman & Rakos, *supra* note 20, at 66, 76; *see also* Margaret Bull Kovera et al., *Jurors' Perceptions of Eyewitness and Hearsay Evidence*, 76 MINN. L. REV. 703, 707, 719 (1992); Peter Miene et al., *Juror Decision Making and the Evaluation of Hearsay Evidence*, 76 MINN. L. REV. 683, 685, 691 (1992).

55. *See generally* Edward J. Imwinkelried, *The Meaning of Probative Value and Prejudice in Federal Rule of Evidence 403: Can Rule 403 Be Used to Resurrect the Common Law of Evidence?*, 41 VAND. L. REV. 879 (1998) (illustrating the inherent difficulties of probative value calculations); Jonathan J. Koehler, *On Conveying the Probative Value of DNA Evidence: Frequencies, Likelihood Ratios, and Error Rates*, 67 U. COLO. L. REV. 859 (1996) (explaining the difficulties of measuring probative value in the scientific evidence context).

currently no prevailing theory of how to appropriately measure various pieces of evidence, nor is there an agreed-upon manner to assess how closely legal decision makers adhere to that measurement.[56]

The second study reported in *Testing Tribe's Triangle: Juries, Hearsay, and Psychological Distance* provides a good example of this problem.[57] In that study, which examined the effects of hearsay within hearsay on jurors' determinations of a defendant's guilt, participants' determinations of the likelihood of the defendant's guilt decreased as the degree of hearsay presented to them increased, which suggests a careful evaluation of the hearsay evidence.[58]

Figure 2: How Jurors Respond to Hearsay Within Hearsay[59]



Yet the slope of the line that connects the means in each experimental condition indicated that there was a drop-off in the amount that jurors discounted when they encountered triple hearsay instead of double hearsay.[60] Although this data suggest that jurors carefully examined the hearsay evidence and made a reasonable judgment regarding the weight to afford it, it is normatively difficult to argue that the jurors weighed the evidence correctly.[61]

In sum, the central rationale for the rule barring hearsay evidence—that unreliable hearsay evidence is a threat to the decisional accuracy of the legal fact finder—is fraught with problems. First, the accuracy of a fact finder's weighing of hearsay evidence is difficult to determine because there is no objective standard for attributing probative value to that evidence, and so it may

56. Indeed, for these reasons, and others, some scientists and scholars have called for an abandonment of using probative value calculations to evaluate the competency of various legal actors. *See, e.g.*, John T. Wixted & Laura Mickes, *The Field of Eyewitness Memory Should Abandon Probative Value and Embrace Receiver Operating Characteristic Analysis*, 7 PERSPECTIVES ON PSYCHOL. SCI. 275 (2012).

57. *See* Sevier, *supra* note 23, at 915–22.

58. *See id.* at 920.

59. This figure originally appeared in *Testing Tribe's Triangle: Juries, Hearsay, and Psychological Distance*. *Id.*

60. *See id.* at 920–21 (reporting the study results).

61. Nonetheless, the clear pattern in the data suggests, at a minimum, reasonable, systematic decision making by mock jurors with respect to hearsay. *Id*.

be impossible to evaluate whether legal fact finders weigh the evidence in the appropriate amount. If so, then the proposition that the hearsay rule safeguards the accuracy of legal verdicts is empirically unfalsifiable, because there is no true way to test that assertion scientifically.

Second, to the extent that we can measure how jurors respond to hearsay, social science evidence strongly suggests that the decisional accuracy rationale for the rule barring hearsay evidence has little empirical basis.[62] The extensive social science research on hearsay evidence suggests, at the very least, that jurors are suspicious of hearsay evidence and view it with measured skepticism.

II. THE SUPERIORITY OF PROCEDURAL JUSTICE

Against this background, I turn to the minority rationale for prohibiting hearsay evidence, and examine whether it is—both conceptually and empirically—the better rationale for the rule. This alternative to the prevailing rationale for the hearsay rule focuses on the dignity interests of litigants who face their accusers in court. This rationale supports prohibiting hearsay not because of its potential to cloud the decision-making process, but because of the perceived unseemliness of the use of such evidence. This rationale has intuitive appeal and, as discussed below, also has empirical support.

At its core, a procedural justice rationale for the hearsay rule stems from the trial of Sir Walter Raleigh, which was discussed at length in *Testing Tribe's Triangle: Juries, Hearsay, and Psychological Distance*.[63] In sum, Raleigh was executed on what scholars now believe to be unreliable hearsay evidence, and his trial sparked outrage amongst the British citizenry.[64] His trial is believed to have been the impetus for, among other legal innovations, the Confrontation Clause of the Sixth Amendment to the United States Constitution.[65]

One important question raised by the trial of Sir Walter Raleigh remains unanswered: what was the source of public outrage with Raleigh's trial? Was the outrage based on the jury's apparent overweighing of the hearsay evidence against Raleigh, or was the outrage based on the Crown's abject refusal to produce as witnesses its hearsay declarants—one of whom had been in the Crown's custody not far from the hall in which Raleigh was tried—despite Raleigh's pleas for the Crown to do so?[66] It is possible that the popular outrage over the verdict in Raleigh's trial was attributable to the latter—and not the former—rationale.

Scholars that subscribe to a procedural justice rationale for the rule barring hearsay evidence argue that there exists a dignity interest in requiring out-of-court accusers to face the accused, which is distinct from the jury's ability (or

62. *See supra* note 54 and accompanying text.
63. *See* Sevier, *supra* note 23, at 881–82 (discussing Raleigh's trial).
64. *Id.*
65. *Id.*
66. *See id.* at 882 (discussing the implications of Raleigh's trial and execution).

inability) to weigh the probative value of out-of-court hearsay statements. Scholars frame this dignity interest in a variety of ways, but at its core, this interest has little to do with decisional accuracy, and instead focuses on notions of fairness and justice in the process used to resolve legal disputes. This focus on process—which social psychologists term procedural justice—might provide a more theoretically coherent and empirically accurate rationale for the hearsay rule, particularly compared to the current, prevailing rationale for excluding hearsay.[67] To that end, this Article explores the concept of procedural justice in greater detail and tests empirically whether it aligns with popular notions of the purpose that the hearsay bar serves in judicial disputes.

In the sections below, I make three claims. First, I argue that, unlike the decisional accuracy rationale for the hearsay rule, dignity and justice concerns over the use of hearsay are supported by legal scholarship as well as experimental social psychology research. Second, I argue that a procedural justice rationale would lead to a more coherent and streamlined hearsay rule compared to how the doctrine currently stands. And finally, I argue—with original empirical evidence—that procedural justice considerations align with laypeople's views about the values that the hearsay rule is designed to protect.

A. THEORETICAL SUPPORT

The minority view expressed by some commentators, rule makers, and scholars[68] invokes principles entirely different from decisional accuracy when justifying the ban on hearsay evidence at trial. Instead, these scholars focus on the effect of the admission of hearsay evidence on the legitimacy of the legal institutions that accept such evidence. This section will first explore the theoretical dignity arguments raised by legal academia, and then will explore the social psychology literature on procedural justice.

1. Legal Theories of Fairness

Although most legal scholarship on the hearsay rule focuses on the role of the rule in safeguarding the accuracy of a legal fact finder's decisions, a few scholars have proposed justifications for the rule barring hearsay evidence that focus instead on notions of fairness and dignity. For example, Professor Eleanor Swift has written that laypeople will likely view verdicts that rely on hearsay evidence as the result of an unfair process—a process in which accusers are not required to look the accused in the eye—which may cause them to discredit the legal system.[69] By this logic, the public outrage over the outcome of the trial of Sir Walter Raleigh did not stem from a belief that Raleigh's tribunal was

67. *See infra* Part II.A and Part II.B for a robust discussion of procedural justice.

68. *See supra* note 22 and accompanying text.

69. *See* Eleanor Swift, *Abolishing the Hearsay Rule*, 75 CALIF. L. REV. 495, 495 (1987) ("[R]eliance on . . . hearsay declarants threatens important values related to the rationality *and fairness* of trial adjudication." (emphasis added)).

hoodwinked by unreliable evidence. Rather, the outrage stemmed from the lack of respect that Raleigh was afforded by not confronting his accuser in court. It is an open question regarding whether the public's view toward hearsay evidence has shifted since Raleigh's trial, but the presence of rules of evidence that bar hearsay evidence[70] as well as a constitutional rule that bars evidence in criminal cases where defendants could not confront their accusers[71] provides some evidence that the same concerns about hearsay evidence that existed during Raleigh's trial still exist today.

Recent Supreme Court decisions examining the intersection between hearsay exceptions and the Confrontation Clause provide additional, if ambivalent, support for the view that the rule against hearsay evidence is derived from concerns about procedural legitimacy and fairness. Although the Court had previously interpreted the Confrontation Clause so that it posed little threat to the admission of hearsay evidence in criminal trials,[72] the Court reversed course in *Crawford v. Washington*.[73] In *Crawford*, the Court evaluated the constitutionality of hearsay statements made to a police officer, which had been corroborated by other evidence in the case.[74] The Court reversed the defendant's conviction and held the hearsay statements inadmissible because the accused had not actually been confronted by his accuser.[75] In reaching this result, the Court argued that the hearsay statements were inadmissible not because they might be unreliable, but because their inclusion in the trial offended a dignity interest to which the defendant is entitled under the Sixth Amendment.[76] Although the Court's decision also included occasional references to the ability of cross-examination to strengthen the reliability of evidence presented to the fact finder,[77] later Supreme Court decisions interpreting *Crawford* continued to focus on the dignity interests of confronting one's accusers at trial when

70. Article VIII of the Federal Rules of Evidence regulates the use of hearsay evidence at trial. *See* FED. R. EVID. 801–807.

71. *See* U.S. CONST. amend. VI (noting that "[i]n all criminal prosecutions, the accused shall enjoy the right . . . to be confronted with the witnesses against him").

72. *See* Ohio v. Roberts, 448 U.S. 56, 66 (1980) ("In sum, when a hearsay declarant is not present for cross-examination at trial, the Confrontation Clause normally requires a showing that he is unavailable. Even then, his statement is admissible only if it bears adequate 'indicia of reliability.' Reliability can be inferred without more in a case where the evidence falls within a firmly rooted hearsay exception. In other cases, the evidence must be excluded, at least absent a showing of particularized guarantees of trustworthiness.").

73. 541 U.S. 36, 69 (2004).

74. *See id.* at 38, 42.

75. *See id.* at 68–69 ("*Roberts* notwithstanding, we decline to mine the record in search of indicia of reliability. Where testimonial statements are at issue, the only indicium of reliability sufficient to satisfy constitutional demands is the one the Constitution actually prescribes: confrontation.").

76. *See id*. at 68 ("Where testimonial evidence is at issue, . . . the Sixth Amendment demands what the common law required: unavailability and a prior opportunity for cross-examination.").

77. *See id*. at 61 (acknowledging the value of testing the reliability of hearsay statements "in the crucible of cross-examination").

deciding whether hearsay evidence is admissible in criminal trials.[78]

Other scholars reach the same conclusion, although their focus is different. Recall that Professor Eleanor Swift expressed concern that if hearsay statements were not generally barred from legal proceedings under a formal set of rules, Federal Rule of Evidence 403—which excludes relevant evidence if its probative value is outweighed by its prejudicial effect[79]—would become the de facto mechanism by which tribunals evaluate hearsay.[80] Professor Roger Park has expressed similar concerns that abolishing the formal rules barring hearsay evidence would give judges too much discretionary power to admit statements that, in their view, are reliable.[81] Thus, evaluating hearsay statements under Rule 403 would likely allow into evidence significantly more hearsay statements than the current rules allow—which raises the concerns expressed by Professor Swift and implied by Professor Park[82]—and would also feed into the perception that legal tribunals arbitrarily decide which hearsay statements to allow into evidence.[83] This perceived arbitrariness would raise additional concerns about the legitimacy of tribunals that allow such evidence to be heard at trial.

---

78. *See* Michigan v. Bryant, 562 U.S. 344, 348–49 (2011) (following *Crawford* and clarifying that the primary purpose of a testimonial hearsay statement is one that does not seek assistance for an ongoing emergency); Giles v. California, 554 U.S. 353, 358–59 (2008) (using actual confrontation as a guide and implying that statements about past emergencies are not testimonial); Davis v. Washington, 547 U.S. 813, 829–30 (2006) (using actual confrontation as its guiding principle and clarifying that a testimonial statement is usually spoken after the events have occurred).

79. "The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence." FED. R. EVID. 403.

80. *See* Swift, *supra* note 69, at 499; *see also* Sevier, *supra* note 23, at 890 nn.47–49 (discussing Professor Swift's contribution to understanding the policy rationales that underlie the hearsay rule).

81. Park, *supra* note 21, at 1060 & n.11 (noting, without explicitly referencing Rule 403, that in such a scheme "judges would have more discretion and stronger control" over hearsay statements that are admitted into evidence and noting that "[t]he fear of unbridled discretion has been one of the bar's primary reasons for opposing these proposals for broader admission of hearsay"). For example, consider the case of *United States v. Bowman*, 302 F.3d 1228 (11th Cir. 2002). Bowman was charged with racketeering, drug offenses, and conspiracy to commit murder when he was the international president of the Outlaws Motorcycle Club. *See id.* at 1231, 1238–39. The Club had a generally unsavory history, and its constitution contained a clause restricting its membership to white men. *See id.* 1238–39. Although Bowman was not charged with any form of discrimination, the district court allowed into evidence an unredacted version of the Club's constitution as proof of other matters in the case, and the defendant was convicted. *See id.* at 1239. The appellate court held that the trial court's ruling was an abuse of its discretion under Federal Rule of Evidence 403, but because of the other evidence against the defendant, the appellate court held that the district court's error was harmless. *Id.* at 1239–40. Decisions regarding hearsay evidence—which would be decided under Rule 403 if the hearsay bar were eliminated—would likely be subject to similar judicial discretion.

82. Professor Park raises additional fairness concerns that underlie the hearsay rule, including the "mischief" of perjury in court (regardless of its effect on accurate fact-finding), the potential for jurors to reach inequitable verdicts by disingenuously relying on hearsay evidence, the potential impacts on the speedy disposition of weak cases, and the cost to parties in obtaining and developing evidence to rebut hearsay statements from out-of-court declarants. *See* Park, *supra* note 21, at 1058–59.

83. *See* Swift, *supra* note 69, at 499; Park, *supra* note 21, at 1060.

2. Social Psychological Support

In addition to legal scholarship that has examined, theoretically, the role that dignity and fairness interests should play in constructing a hearsay rule that attains popular legitimacy, there also is a plethora of social psychological evidence that suggests the prohibition on hearsay evidence serves an important—and real—dignity function. This research on the psychological phenomenon of procedural justice is outlined below.

Traditional law and economics scholars have hypothesized that people seek to maximize their self-interest and pursue opportunities that bring about the greatest possible material advantage.[84] If this were always true, we would expect that people prefer legal systems that provide them with the best economic outcomes.[85] Although research suggests that outcomes do matter with respect to people's procedural preferences,[86] the full psychological story is more complex. In conferring legitimacy onto a legal decision, people are remarkably sensitive to whether the process for reaching that decision was fair.[87] This phenomenon, known as procedural justice, predicts that "people's reactions to their experiences with legal authorities are strongly shaped by their subjective evaluations of the justice of the procedures used to resolve their case."[88] In short, independent of outcomes, the manner in which a legal dispute is decided can predict people's preferences for certain legal procedures, perceptions of the decision maker's legitimacy, and their willingness to abide by the decision maker's judgments.

Researchers have identified several procedural factors that influence people's perceptions of the legitimacy of a decision-making body: the decision maker's neutrality (that is, that the decision is based on rules and facts instead of the decision maker's intuition), the degree of respect and dignity that the decision maker confers on the parties, the amount of voice and control that the parties

84. *See generally* LYNN A. STOUT, CULTIVATING CONSCIENCE: HOW GOOD LAWS MAKE GOOD PEOPLE (2011) (discussing empirical evidence on the role of an individual's conscience in everyday decision making and arguing that people frequently make decisions in daily life that traditional law and economics scholars would classify as irrational). *See also* JOHN W. THIBAUT & HAROLD H. KELLEY, THE SOCIAL PSYCHOLOGY OF GROUPS (1959) (noting the traditional law and economics rationality arguments and examining them in the context of social interactions involving group norms, interdependence, and rewards and punishments).

85. *See generally* LIND & TYLER, *supra* note 24 (examining the conditions under which laypeople will legitimize the judgments of a decision maker and finding that several factors, including the perceived fairness of the process used by the decision maker, affects the willingness of laypeople to legitimize the decision maker's judgments).

86. Tom R. Tyler & E. Allan Lind, *Procedural Justice*, *in* HANDBOOK OF JUSTICE RESEARCH IN LAW 65, 71 (Joseph Sanders & V. Lee Hamilton eds., 2001).

87. *See generally* Tom Tyler & David Markell, *The Public Regulation of Land-Use Decisions: Criteria for Evaluating Alternative Procedures*, 7 J. EMP. LEGAL STUD. 538, 569–70 (2010). *See also* E. Allan Lind, Robin I. Lissak & Donald E. Conlon, *Decision Control and Process Control Effects on Procedural Fairness Judgments*, 13 J. APPLIED SOC. PSYCHOL. 338, 348 (1983) ("The results of the present study suggest that there is a process control effect on fairness *independent* of the decision control accorded those affected by the procedure's outcome.").

88. Tyler & Markell, *supra* note 87, at 541 (internal citations omitted).

have over the legal dispute, and the degree to which parties can trust the decision maker's motive to be fair.[89] These factors manifest themselves in the laboratory and outside the laboratory[90] and in both criminal and civil disputes.[91]

Procedural justice is a robust phenomenon and its implications are far-reaching. In legal adjudication, perceptions of fair process can confer legitimacy on legal actors, including judges[92] and juries.[93] Procedural justice can also affect perceptions of legitimacy in alternative dispute resolution, including mediation and arbitration, as well as the legitimacy of the decision makers in those paradigms.[94]

Not only do perceptions of fair treatment influence people's preferences in the courtroom, they also influence people's preferences in the legal system outside the courtroom and in nonlegal contexts.[95] For example, affording those who have been stopped by the police an explanation for their detainment and a chance to explain themselves to a law enforcement officer increases people's perceptions of law enforcement legitimacy.[96] Moreover, perceptions of procedural fairness can increase perceptions of legitimacy (and a willingness to abide by decisions) in nonlegal relationships that involve power dynamics, including superior-subordinate relationships in the workplace[97] and in the family.[98]

Psychological researchers have sought to understand the reasons why laypeople care about the process by which legal adjudications are made. In a famous set of studies expounding on the "value-expressive" function of procedural justice,[99] Professor Tom Tyler and others found that, by allowing litigants

89. *See* Tom R. Tyler, *Social Justice: Outcome and Procedure*, 35 INT'L J. PSYCHOL. 117, 117 (2000).

90. *See* Rebecca Hollander-Blumoff, *The Psychology of Procedural Justice in the Federal Courts*, 63 HASTINGS L.J. 127, 135 (2011).

91. *See* E. ALLAN LIND, ARBITRATING HIGH-STAKES CASES: AN EVALUATION OF COURT-ANNEXED ARBITRATION IN A UNITED STATES DISTRICT COURT 13–14 (1990), http://www.rand.org/content/dam/rand/pubs/reports/2007/R3809.pdf (evaluating these factors in the context of civil disputes that are the subject of court-annexed arbitration); *see also* Justin Sevier, *The Truth-Justice Tradeoff: Perceptions of Decisional Accuracy and Procedural Justice in Adversarial and Inquisitorial Legal Systems*, 20 PSYCHOL. PUB. POL'Y & L. 212, 214–15 (2014) (finding no difference between criminal and civil settings with respect to participants' perceptions of procedural justice in a mock trial vignette study).

92. *See* TYLER, *supra* note 24, at 19, 30–33.

93. *See* Robert J. MacCoun & Tom R. Tyler, *The Basis of Citizens' Perceptions of the Criminal Jury: Procedural Fairness, Accuracy, and Efficiency*, 12 LAW & HUM. BEHAV. 333, 350 (1988).

94. *See* Rebecca Hollander-Blumoff & Tom R. Tyler, *Procedural Justice in Negotiation: Procedural Fairness, Outcome Acceptance, and Integrative Potential*, 33 LAW & SOC. INQUIRY 473, 478–79 (2008) (discussing negotiation); *see also* LIND, *supra* note 91 (discussing arbitration).

95. *See* Hollander-Blumoff, *supra* note 90, at 133.

96. *See* Tom R. Tyler & Robert Folger, *Distributional and Procedural Aspects of Satisfaction with Citizen-Police Encounters*, 1 BASIC & APPLIED SOC. PSYCHOL. 281 (1980).

97. *See* Tom R. Tyler, *Promoting Employee Policy Adherence and Rule Following in Work Settings: The Value of Self-Regulatory Approaches*, 70 BROOK. L. REV. 1287, 1288 (2005).

98. *See* Shelly Jackson & Mark Fondacaro, *Procedural Justice in Resolving Family Conflict: Implications for Youth Violence Prevention*, 21 LAW & POL'Y 101, 118–20 (1999).

99. Tom R. Tyler, Kenneth A. Rasinski & Nancy Spodick, *Influence of Voice on Satisfaction with Leaders: Exploring the Meaning of Process Control*, 48 J. PERSONALITY & SOC. PSYCHOL. 72, 73 (1985). In a related study, researchers provided participants with a set of demanding, time-sensitive tasks to

an opportunity to present their side of the dispute to the decision maker,[100] by affording them respect as they do so,[101] and by doing so in an unbiased manner,[102] litigants felt valued and respected by the decision maker as members of society.[103]

Thus, the mere ability of participants to express themselves to the tribunal in these studies was found to predict participants' feelings of dignity, respect, and the fairness of the process for adjudicating the dispute. In the hearsay context, we can imagine similar dignity interests implicated. These studies suggest that laypeople focus strongly not only on substantive outcomes, but also on the process for obtaining those outcomes, when they evaluate their satisfaction with a legal proceeding. The empirical evidence therefore implies that, to the extent that laypeople dislike hearsay evidence, their concerns over this type of evidence may relate to the perceived fairness of the way hearsay is used in court, rather than the substantive effect it may have on the tribunal's decision.[104]

---

complete, and participants were assigned to one of three conditions: (1) a forum in which they were not allowed to present their opinions about the tasks to the decision maker; (2) a forum in which they were allowed to do so; and (3) a forum in which they were allowed to present their opinions after the task was completed, and were explicitly told that their opinions would have no effect on the authority figure's decisions. *See* E. Allan Lind, Ruth Kanfer & P. Christopher Earley, *Voice, Control, and Procedural Justice: Instrumental and Noninstrumental Concerns in Fairness Judgments*, 59 J. PERSONALITY & SOC. PSYCHOL. 952, 954–55 (1990). When asked how satisfied they were with the proceeding, participants who were not allowed to present their opinions were the least satisfied and participants who were able to present their opinions were the most satisfied. *See id*. at 955–56. But the participants who could present their opinions after-the-fact—even though they knew it would not affect the outcome—were significantly more satisfied than those who could not present their side of the case, even though they received the same substantive outcome. *See id.*

100. *See* Tom R. Tyler, *Conditions Leading to Value-Expressive Effects in Judgments of Procedural Justice: A Test of Four Models*, 52 J. PERSONALITY & SOC. PSYCHOL. 333, 341 (1987) (describing "consideration of views" as an important mediator of "voice effects" in citizens' perceptions of procedural justice).

101. *See id*. at 343 ("Studies in the legal, managerial, and political arenas have found that citizens place great value on being treated politely and having respect shown for their rights. These are aspects of the interaction that are directed at reinforcing citizens' positive self-image and sense of personal worth . . . ."); Tyler & Markell, *supra* note 87, at 548 (discussing the role of trust and respect in laypeople's perceptions of procedural justice).

102. *See* Tom R. Tyler, *The Psychology of Procedural Justice: A Test of the Group-Value Model*, 57 J. PERSONALITY & SOC. PSYCHOL. 830, 831 (1989) ("In any particular situation people will be concerned with having an unbiased decision maker who is honest and who uses appropriate factual information to make decisions.").

103. *See* Tyler, *supra* note 100, at 343 (describing the impact of voice effects on a citizen's self-esteem, self-image, and perception of personal worth); *see also* Tom R. Tyler & E. Allan Lind, *A Relational Model of Authority in Groups*, 25 ADVANCES IN EXPERIMENTAL SOC. PSYCHOL. 115, 165 (1992) (noting the effects of "standing, trust, and neutrality" on the willingness of people to legitimize authority figures and explaining that "the source of information about these factors is the social process and the procedures that the authority uses"). *See generally* Tyler, *supra* note 102 (reporting an empirical test of the "group-value" model of procedural justice, which hypothesizes that concerns over the fairness of adjudicative processes implicates social psychological factors that influence self-esteem).

104. For a nonempirical perspective on procedural legitimacy, see generally Solum, *supra* note 24.

B. DOCTRINAL COHERENCE

A procedural justice rationale for the rule barring hearsay is supported by empirical social psychology research. Armed with this understanding, evidence policymakers can use a procedural justice rationale for the hearsay rule to make the rule more conceptually coherent. Particularly, a procedural justice rationale for the hearsay rule would eliminate many of the empirically dubious exceptions to the rule, which rely on empirically untested notions regarding the reliability of such evidence. It would also eliminate the current disparity between the role of the Confrontation Clause in cases involving hearsay in civil and criminal disputes. These issues are discussed in detail below.

The general proposition that second-hand evidence should be barred from dispute resolution proceedings because the evidence has not been cross-examined appears, initially, as a straightforward rule. The difficulty, however, is that this straightforward principle contains nearly thirty exceptions, which are found primarily in Federal Rules of Evidence 801, 803, and 804.[105] Moreover, the rationales for these exceptions are wildly different: the Rule 803 exceptions focus on reliability and decisional accuracy, whereas Rule 804 exceptions focus on fairness and equity.[106]

Many of the exceptions to the rule barring hearsay evidence under Federal Rule of Evidence 803 could be streamlined dramatically by focusing on procedural justice as the rationale for disallowing hearsay evidence, instead of decisional accuracy. Rule 803's reliability exceptions, as scholars have noted, are based on untested notions about human cognition and behavior.[107] It is not obvious that an utterance made while a speaker is excited is likely to be truthful,

---

105. *See* FED. R. EVID. 801(d)(1)–(2), 803(1)–(23), & 804(1)–(6).

106. *Compare* FED. R. EVID. 803 advisory committee's note (stating, while providing no empirical evidence in support, that "[t]he present rule proceeds upon the theory that under appropriate circumstances a hearsay statement may possess circumstantial guarantees of trustworthiness sufficient to justify nonproduction of the declarant in person at the trial even though he may be available"), *and* FED. R. EVID. 803(b)(1) advisory committee's note (explaining the rationale for the excited utterance exception as "simply that circumstances may produce a condition of excitement which temporarily stills the capacity of reflection and produces utterances free of conscious fabrication"), *with* FED. R. EVID. 804(b) advisory committee's note ("Rule 803 . . . is based upon the assumption that a hearsay statement falling within one of its exceptions possesses qualities which justify the conclusion that whether the declarant is available or unavailable is not a relevant factor in determining admissibility. The instant [R]ule [804] proceeds upon a different theory: hearsay which admittedly is not equal in quality to testimony of the declarant on the stand may nevertheless be admitted if the declarant is unavailable and if his statement meets a specified standard. . . . [H]earsay, if of the specified quality, is preferred over complete loss of the evidence of the declarant.").

107. *See* Maithilee K. Pathak & William C. Thompson, *From Child to Witness to Jury Effects of Suggestion on the Transmission and Evaluation of Hearsay*, 5 PSYCHOL. PUB. POL'Y & L. 372, 373 (1999) ("However, these fears are primarily grounded in a psychological intuition: Courts have traditionally assumed, without much empirical basis, that jurors do not appreciate fully the unreliability of hearsay testimony . . . ."); Justin Sevier, *Omission Suspicion: Juries, Hearsay, and Attorneys' Strategic Choices*, 40 FLA. ST. U. L. REV. 1, 7 (2012) ("For example, the law of evidence is premised largely on codified but empirically untested folk wisdom regarding how legal actors, like jurors, behave.").

as Rule 803(2) supposes,[108] or that people are always honest with their health-care providers, as Rule 803(4) supposes.[109]

These empirically dubious behavioral assumptions can cause not only inaccurate legal decisions, but also moral hazards with respect to litigants' pretrial actions in which savvy litigants could decide to strategically lie to their doctors on the expectation that the otherwise inadmissible lie would be admissible under a hearsay exception. A procedural justice rationale for the hearsay rule would remove these empirically dubious exceptions to the rule, while the other reliability exceptions that focus on the integrity of books, records, or other documents[110] could remain if policymakers believe that this evidence does not offend the procedural dignity of the party to whom it is used against.[111]

A procedural justice rationale for the hearsay rule would also add doctrinal coherence with respect to the manner in which the hearsay rule is applied in criminal and civil trials. The hearsay rule currently intersects with the Sixth Amendment Confrontation Clause only in criminal cases.[112] Under the Supreme Court's current Confrontation Clause jurisprudence, the same out-of-court statement (that was excluded in a criminal trial) is admissible in a civil proceeding against the defendant, provided that the statement fits into one of the hearsay exceptions listed in Rule 803 or 804. The rationales provided for this differential admissibility of testimonial hearsay—which center on increased protections for criminal defendants relative to civil defendants—are viewed by many scholars as unconvincing,[113] and a procedural justice rationale for the hearsay rule would remedy this inequality by excluding testimonial hearsay from both tribunals.

The U.S. Supreme Court's Confrontation Clause jurisprudence, of course, represents a floor below which an individual's rights cannot fall.[114] As a matter of policy, states and the federal government are free to expand on those rights and provide greater protections to individuals. If the hearsay rule is premised on the notion that, no matter the tribunal, the accused has a dignity interest in facing her accuser, the same principles that underlie the Supreme Court's

---

108. *See* FED. R. EVID. 803(2) advisory committee's note.

109. *See* FED. R. EVID. 803(4) advisory committee's note. Indeed, the limited empirical evidence on this topic suggests the opposite. *See, e.g.*, Karen Ravn, *Body of Lies: Patients Aren't 100% Honest with Doctors*, L.A. TIMES (June 8, 2009), http://articles.latimes.com/2009/jun/08/health/he-lying8.

110. *See* FED. R. EVID. 803(6)–(18).

111. Indeed, while streamlining empirically dubious exceptions based on the so-called reliability of certain hearsay statements, a procedural justice paradigm could allow for many of the fairness-based exceptions found in Federal Rule of Evidence 804 to remain intact.

112. *See* Crawford v. Washington, 541 U.S. 36, 42 (2004).

113. *See, e.g.*, Laird C. Kirkpatrick, *Confrontation and Hearsay: Exemptions from the Constitutional Unavailability Requirement*, 70 MINN. L. REV. 665 (1986); Park, *supra* note 22; Swift, *supra* note 22.

114. *See* Ilya Somin, *A Floor, Not a Ceiling: Federalism and Remedies for Violations of Constitutional Rights in* Danforth v. Minnesota, 102 NW. U. L. REV. COLLOQUY 365, 365 (2008) ("Few doubt that states can provide greater protection for individual rights under state constitutions than is available under the Supreme Court's interpretation of the Federal Constitution.").

reasoning in its Confrontation Clause jurisprudence could be applicable to civil trials as well.

A procedural justice rationale for excluding hearsay evidence would thus have dual advantages. It would eliminate many of the reliability exceptions that lack scientific support while potentially allowing many of the exceptions designed to promote equity concerns to remain. It would also harmonize the disparate treatment of testimonial hearsay in criminal and civil trials. A procedural justice rationale will also likely lead to greater popular acceptance and perceived legitimacy of the hearsay rule. Two experiments designed to test this proposition are discussed below.

C. POPULAR LEGITIMACY

A procedural justice rationale for the hearsay rule has empirical support from social science research and would lead to a more coherent and streamlined doctrine. Past empirical research also suggests that basing the hearsay rule on procedural justice concerns may lead to a rule that the public views as more legitimate.[115] Indeed, legal scholars have frequently argued that substantive legal doctrines should align more closely with popular conceptions of those doctrines.[116]

No one has examined empirically whether laypeople support or oppose the use of hearsay evidence. Moreover, no one has examined the reasons for which laypeople may oppose hearsay evidence, even though such data could affect not only the direction of future hearsay scholarship, but also the direction of hearsay policy itself. Two original studies reported here provide empirical evidence that laypeople conceive the hearsay rule as one that protects the dignity interests of litigants and not the accuracy of legal decisions.[117]

The first study examines whether laypeople's satisfaction with a dispute resolution proceeding is a function of the degree to which the proceeding includes hearsay evidence. If, consistent with psychological research on proce-

---

115. *See generally* THIBAUT & WALKER, *supra* note 24 (premising their book on the notion that concerns about the process by which disputes are litigated have an effect on people's attitudes toward the decision maker); TYLER, *supra* note 24 (same).

116. *See, e.g.*, Kenworthey Bilz, *Dirty Hands or Deterrence? An Experimental Examination of the Exclusionary Rule*, 9 J. EMPIRICAL LEGAL STUD. 149, 154–66 (2012) (empirically testing two competing rationales for the exclusionary rule, determining that laypeople support the integrity justification, and arguing that the Supreme Court's exclusionary rule jurisprudence should align with this justification); Paul H. Robinson, Michael T. Cahill & Daniel M. Bartels, *Competing Theories of Blackmail: An Empirical Research Critique of Criminal Law Theory*, 89 TEX. L. REV. 291, 313–34, 346 (2010) (examining competing theories of blackmail to determine which has the most popular support and recommending that the law align with the public's understanding of the crime).

117. These studies are substantially in the form of a controlled laboratory experiment, in which the experimental variable of interest is manipulated, whereas all other aspects of the experiment remain the same in every experimental condition. The controlled laboratory experiment is the best way to make statements of causality with respect to the manipulated variables, but experiments also come with significant tradeoffs with respect to ecological validity. For a robust discussion of these issues, see *infra* Part V.B.

dural justice, participants' discomfort with hearsay evidence is a function of the dignity interest of the defendant—and not the effect of hearsay evidence on the risk of inaccurate verdicts—this provides empirical support for the proposition that a streamlined hearsay rule that focuses on procedural justice, and not decisional accuracy, will receive greater popular legitimacy insofar as it more closely aligns with popular conceptions of the purpose of the hearsay rule.

The second study extends the empirical analysis reported in the first study by examining participants' spontaneous responses to two trials that either contain, or do not contain, hearsay evidence. If participants are dissatisfied with hearsay evidence, and their direct responses to questions posed to them specifically reference procedural justice concerns instead of decisional accuracy concerns, the second study provides even greater evidence that a procedural justice rationale for the hearsay rule is likely to receive greater popular legitimacy.

III. STUDY 1

The first study reported in this Article tests the assertion that laypeople will be less satisfied with a dispute resolution proceeding that contains hearsay evidence compared to a proceeding in which the evidence is presented as direct testimony by a live witness. Moreover, it will assess whether any dissatisfaction laypeople may experience with hearsay evidence is because hearsay evidence raises the risk of inaccurate decisions or because hearsay evidence is offensive to the dignity interest of the defendant. The study will also examine whether this hypothesized discomfort—and the psychological mechanism that underlies it—exists in different types of dispute resolution proceedings.

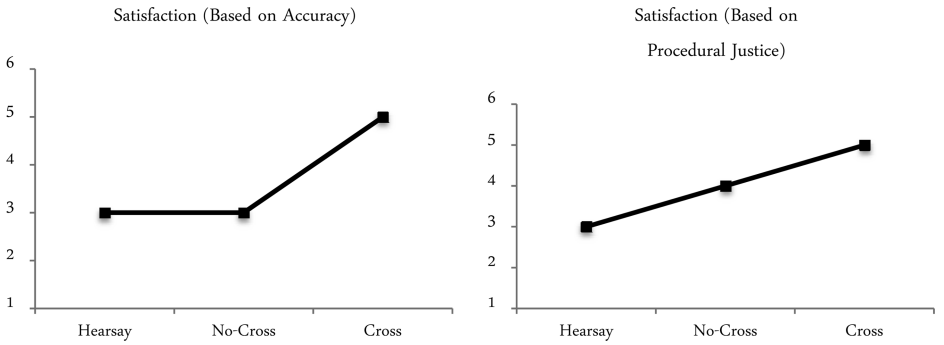
If participants dislike hearsay evidence, as suggested by the empirical research to date,[118] we would expect lower ratings of satisfaction for trials that contain hearsay evidence compared to those that contain a live witness who is subjected to cross-examination. But Study 1 contains an additional version of a trial, in which the witness appears live, but the witness is not subjected to cross-examination. Thus, the only difference between this version of the trial and the version that contains hearsay evidence is the in-court presence of the witness; the witness otherwise provides exactly the same information to the tribunal in both versions of the trial. If participants believe that the hearsay rule prevents erroneous verdicts, we would expect them to be just as dissatisfied with this additional condition, because the witness's mere presence—without being exposed to the crucible of cross-examination—should not advance this accuracy goal.[119] But if participants believe that the hearsay rule serves a

118. *See supra* note 54 and accompanying text (discussing experiments that suggest that jurors give hearsay evidence less probative weight than they give other types of evidence).

119. Astute readers may wonder, however, whether participants in this experimental condition would simply trust that the court is exercising all of its procedural safeguards when a witness appears live. In other words, even though these particular witnesses are not subject to cross-examination, perhaps the participants would believe that there is a good reason for the lack of cross-examination (and so their

dignity interest—by forcing accusers to look the accused in the proverbial eye—then we might expect jurors to be more satisfied with this additional condition, even though the witness was not subject to cross-examination. The figure below illustrates the different predictions based on these differing policy goals.

Figure 3: Hypothesized Satisfaction Reported in the Mock Trial by Experimental Condition



It is possible, however, that participants believe that the hearsay rule safeguards decisional accuracy, and that the mere presence of the witness (without cross-examination) still furthers this goal, because the tribunal may be able to assess the witness's demeanor.[120] To distinguish such a rationale from participants who may believe that the witness's presence serves a dignity function, participants will also be asked the degree to which they believe the trial is fair to the defendant, and the degree to which they believe the tribunal will be able to reach an accurate verdict. Participants' responses can then be compared to their satisfaction ratings to determine which rationale better accounts for those ratings. For example, if a concern over decisional accuracy drives participants' judgments of satisfaction with the trial, we would expect results that approximate the left side of Figure 3, in which the mere production of the witness has little to no effect on the accuracy of the tribunal's decision. If, however, a concern for procedural justice and the dignity interest of the defendant drives their satisfaction judgments, we would expect a pattern of results similar to

accuracy concerns are allayed because they trust the court is doing its job correctly). This is a credible alternative hypothesis for the no-cross experimental condition, and so we might expect that perceptions of accuracy in this condition would be similar to the cross condition instead of the hearsay condition in Figure 3.

120. Such beliefs are misplaced. *See, e.g.*, Paul Ekman & Maureen O'Sullivan, *Who Can Catch a Liar?*, 46 AM. PSYCHOLOGIST 913, 920 (1991) (arguing, with empirical evidence, that laypeople are unskilled at assessing the credibility of others); Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 787 (2001) (finding that even professional legal decision makers are subject to mistakes of reasoning and impression formation).

those on the right side of Figure 3. Specifically, we would expect that the mere production of the witness will increase participants' satisfaction with the trial.[121]

A. PARTICIPANTS

Three hundred twenty-one participants were recruited for an online study using the web-recruiting service Amazon Mechanical Turk (MTurk), which has been shown empirically to be an inexpensive way to collect quality data from persons who are representative of the general Internet-using population.[122] The sample of participants in this study was 45.80% female, 72.50% Caucasian, and averaged 35.73 years of age (with a standard deviation of 12.21 years). Approximately 89.40% of participants in the sample had completed at least a college degree, and the median household income of the sample was between $40,000 and $49,999.

B. PROCEDURE AND MEASURES

Participants were randomly assigned to one of six experimental conditions in the study. Specifically, participants were assigned to an experimental condition that contained one type of dispute resolution procedure and one type of key evidence. The dispute resolution mechanism to which participants were exposed included either a criminal trial or a quasi-administrative hearing. The key evidence against the accused consisted of (1) a hearsay statement, (2) a live witness who was not subject to cross-examination, or (3) a live witness who was subject to cross-examination. The dispute resolution mechanism to which participants were exposed and the primary evidence against the accused were randomly assigned to each participant.[123]

Participants then read a set of written materials. Half the participants read a vignette depicting a criminal conspiracy trial, and half the participants read a vignette depicting an internal dispute resolution procedure at a major corporation. In the criminal conspiracy trial, the defendant was accused of conspiring with a coworker to assassinate a state senator. The State alleged that the defendant, who had a vendetta against the senator, approached the coworker at the senator's town hall meeting and eventually formed a plan to purchase a rifle,

---

121. If neither decisional accuracy nor procedural justice drives participants' satisfaction with the trial, we would expect no relationship between each rationale and participants' satisfaction judgments. If both concerns drive their satisfaction judgments, we should see the same pattern of results for satisfaction, decisional accuracy, and fairness across the different experimental conditions.

122. *See, e.g.*, Adam J. Berinsky, Gregory A. Huber & Gabriel S. Lenz, *Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk*, 20 POL. ANALYSIS 351, 366 (2012); Michael Buhrmester, Tracy Kwang & Samuel D. Gosling, *Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?*, 6 PERSPECTIVES ON PSYCHOL. SCI. 3, 3, 5 (2011); Winter Mason & Siddharth Suri, *Conducting Behavioral Research on Amazon's Mechanical Turk*, 44 BEHAV. RES. METHODS 1, 2–4 (2012).

123. The number of experimental conditions can be calculated by multiplying together the number of hearing types (two) by the number of evidence conditions (three), which totals six different conditions.

which the defendant would use to shoot the senator. The defendant was then tried for conspiracy to commit murder.

In the corporation's internal dispute resolution proceeding, the defendant was accused of conspiring with a coworker to get his boss fired, whom he strongly disliked. The company alleged that the defendant approached the coworker and formed a plan in which the defendant would purchase drugs and plant them in his boss's office. The authorities would then be called to investigate. In all variations of the dispute, the coworker confessed to his involvement in the alleged conspiracy and implicated the defendant. The defendant was then the subject of an internal dispute resolution proceeding, in which he would be terminated from the company if he were found to have committed the acts of which he was accused.

In both disputes, information from the coworker consisted of one of three forms. The coworker either: (1) provided a lengthy written confession of the conspiracy, read into the record by the State (or the corporation), in which he implicated the defendant; (2) testified live in the hearing to these same facts and was not cross-examined; or (3) testified live in the hearing to these same facts but was cross-examined. Every other aspect of each dispute remained exactly the same, regardless of the experimental condition to which the participant was assigned. After reading closing arguments by each party, participants were told that the fact finder began its deliberation.

Participants were not informed of the fact finder's verdict and were instead asked several questions regarding what they had read. First, I asked questions designed to gauge the degree to which participants had paid attention to the dispute resolution proceeding to which they had been exposed. For example, participants were asked questions regarding the type of evidence that was used against the defendant, the nature of the charges against him, and the form of the evidence against him. Participants could not continue the experiment unless they answered these questions correctly.[124]

Next, participants were asked a series of questions designed to assess their general attitudes regarding the dispute resolution procedure. They were asked to rate, on a scale ranging from 1 (very unsatisfied) to 7 (very satisfied), how satisfied they were with the dispute resolution proceeding. They were also asked to rate, on similar seven-point scales, the degree to which the dispute resolution proceeding would produce an accurate verdict and the extent to which it would

124. At this time, there is no way to ascertain on MTurk how many participants answered the comprehension questions correctly on their first attempt. There is, however, circumstantial evidence to suggest that nearly all participants answered the comprehension questions accurately on their first attempt, and that most participants simply did not guess randomly until they were able to advance in the study. The survey software does collect data with respect to the amount of time that participants spent on the page of the study that contained the comprehension questions, and an examination of that data suggests that most participants spent roughly the same amount of time answering the questions. Because the comprehension items consisted of several multiple-choice questions, random guessing should have caused those participants to spend considerably more time answering the questions correctly.

produce a fair or just verdict. Afterward, participants answered how satisfied they were with the dispute resolution procedure in a dichotomous form: that they were satisfied or unsatisfied.

I then assessed each participant's impressions of the evidence presented to them. For example, they were asked to rate how satisfied they were with the coworker's information (either in its confession form or in its live-testimony form), as well as how fair or just the inclusion of that testimony was and how likely it was to help the court reach an accurate verdict. Each of these questions required participants to answer on a seven-point scale, with more negative views anchored at 1 and more positive views anchored at 7.

After completing these questions, participants were asked to provide demographic information, including their age, race, income, level of education completed, and their familiarity with the legal system. Participants were then debriefed with respect to the aims of the study, and the experiment was concluded.

C. RESULTS

This analysis examines the effect of the type of hearing and the hearsay manipulation on participants' attitudes about the trial. I conducted an analysis of variance (ANOVA) to determine whether the type of hearing, the hearsay manipulation, or a combination of these two variables affected participants' satisfaction with the trial, perceptions of its fairness, and perceptions of its ability to render a factually accurate verdict.[125]

1. Satisfaction

I first analyzed the data provided by the study participants with respect to their scaled ratings of satisfaction. As predicted, the ANOVA revealed two

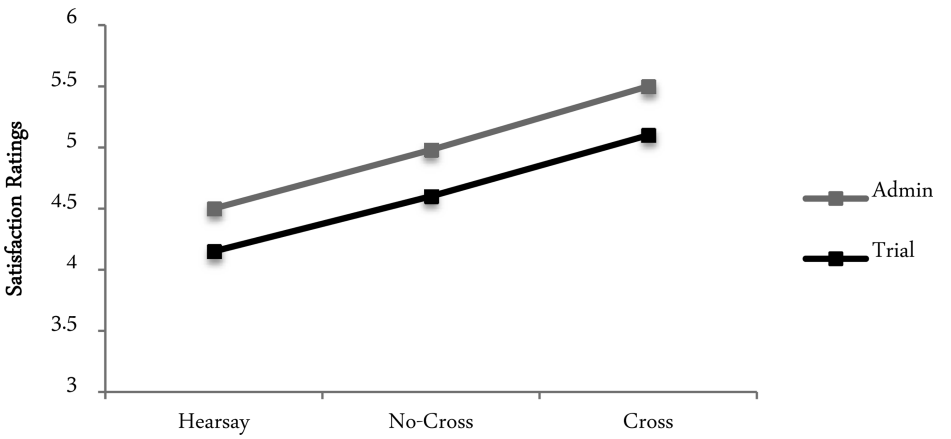
125. Specifically, I conducted a two (hearing type: administrative vs. criminal) x three (type of hearsay: confession vs. live witness with no cross-examination vs. live witness with cross-examination) ANOVA on participants' satisfaction with the trial. I then conducted the same ANOVA on participants' perceptions of the fairness of the trial and on their perceptions of the ability of the court to render an accurate factual decision.

An analysis of variance provides a statistical test of whether the means of several groups are equal. ANOVA results are represented by an F-statistic, and the sizes of the effects are represented by η^2_p. Means are denoted by the letter "M" and standard deviations are denoted by the letters "SD." *See* ROBERT M. LAWLESS ET AL., EMPIRICAL METHODS IN LAW 277–85 (2010) (explaining empirical research methodologies and statistical techniques).

Differences are denoted as "statistically significant" in this Article if the statistical tests indicate that the likelihood that the difference observed would occur by chance is 5% or less (as indicated by the p-value as $p < 0.05$). A difference is "marginally significant" if the likelihood of seeing such a difference by chance is greater than 5% but less than 10%. Jennifer K. Robbennolt, *Apologies and Legal Settlement: An Empirical Examination*, 102 MICH. L. REV. 460, 485 n.117 (2003) (citing BARBARA G. TABACHNICK & LINDA S. FIDELL, USING MULTIVARIATE STATISTICS (2d ed. 1989)). Planned comparisons were accompanied by the Tukey HSD Test to stabilize the "experimentwise error rate" and avoid false positives. *See, e.g.*, James Jaccard, Michael A. Becker & Gregory Wood, *Pairwise Multiple Comparison Procedures: A Review*, 96 PSYCHOL. BULL. 589, 594–95 (1984) (discussing several techniques, including the Tukey technique, for controlling Type I error when making multiple comparisons among groups).

statistically significant findings. First, the analysis revealed a significant finding based on the type of hearing, such that participants were more satisfied, on average, with the quasi-administrative hearing than with the criminal trial.[126] Second (and more importantly), the analysis revealed that the type of hearsay evidence affected participants' satisfaction with the trial. Specifically, participants were the least satisfied when the evidence against the defendant was in the form of the written confession, more satisfied when that evidence was presented in the form of a live witness who was not cross-examined, and the most satisfied when that evidence was presented in the form of a live witness subject to cross-examination.[127]

Figure 4: Ratings of Satisfaction with the Tribunal in an Administrative and Trial Setting



To determine whether participants' satisfaction with the criminal trial or the administrative hearing was the result of a linear relationship among the hearsay confession, live witness (with no cross-examination), and live witness (with cross-examination) versions of the procedure to which they were exposed—and which is suggested by the graphs in the figure above—I employed a statistical

126. $M_{\text{criminal}} = 4.48$ ($SD = 1.90$), $M_{\text{administrative}} = 5.02$ ($SD = 1.86$); $F(1, 315) = 8.13$, $p = .005$, $\eta^2_p = .03$.

127. $M_{\text{hearsay}} = 4.33$ ($SD = 1.94$), $M_{\text{no-cross}} = 4.73$ ($SD = 1.77$); $M_{\text{cross}} = 5.15$ ($SD = 1.93$); $F(2, 315) = 5.90$, $p = .003$, $\eta^2_p = .04$. This pattern occurred regardless of the type of hearing (the relative averages were simply lower across the board in all three criminal trial conditions compared to the three administrative trial conditions). Further, there was no interaction between the type of hearing to which participants were exposed and the type of hearsay evidence presented, $F(2, 315) = 0.55$, $p = .576$.

Although the difference in satisfaction between the administrative hearing and the trial condition is intriguing, the remaining analyses—examining the role of decisional accuracy and procedural fairness in explaining participants' ratings of satisfaction—will average the administrative hearing and trial conditions together for simplicity. Full results separated by hearing condition are on file with the author.

technique called a contrast analysis.[128] As predicted, the contrast analysis revealed a statistically significant linear pattern consistent with the study's hypothesis. Specifically, as the evidence presented to the participants changed from an out-of-court hearsay confession, to the same confession repeated in court by that statement's author, to a live cross-examination of the confessor, participants became more satisfied with both the criminal trial and the administrative hearing.[129]

I found similar results when I examined participants' dichotomous ratings of satisfaction, in which participants selected whether they were generally satisfied or generally dissatisfied with the dispute resolution procedure to which they were exposed.[130] The results revealed a statistically significant effect of the type of hearsay exposed to participants. Compared to the hearsay confession condition, participants exposed to the live witness (with no cross-examination) were two times more likely to express satisfaction, and participants exposed to the live witness who was cross-examined were three times more likely to express satisfaction than were participants who were exposed to the hearsay confession.

2. Fairness

I conducted the same analysis with respect to participants' perceptions of the perceived fairness of the hearing to which they were exposed. Again (and as predicted), the ANOVA revealed that the hearsay condition affected participants' judgments of fairness, such that participants' perceptions of the fairness to the accused were lowest in the hearsay confession condition, higher in the live witness (no cross-examination) condition, and highest in the live witness (with cross-examination) condition.[131] As predicted, a contrast analysis revealed a meaningful linear pattern in the perceptions of fairness for participants exposed to the administrative hearing and those exposed to the criminal trial; as the procedure allowed the defendant more control over the trial, participants' perceptions of the fairness of the procedure gradually increased.[132]

---

128. A contrast analysis determines whether a statistically significant linear pattern exists within the data. Whereas an ANOVA allows researchers to determine whether any of the means for several different groups are different, a contrast analysis allows researchers to test more specific hypotheses, for example, whether the means show a specific polynomial pattern, such as a linear, cubic, or quadratic function. In sum, a contrast analysis tests a specific question about the pattern of results revealed in an ANOVA. *See, e.g.*, *Contrast Analysis*, *in* 1 ENCYCLOPEDIA OF RESEARCH DESIGN 243, 243–44 (Neil J. Salkind ed., 2010).

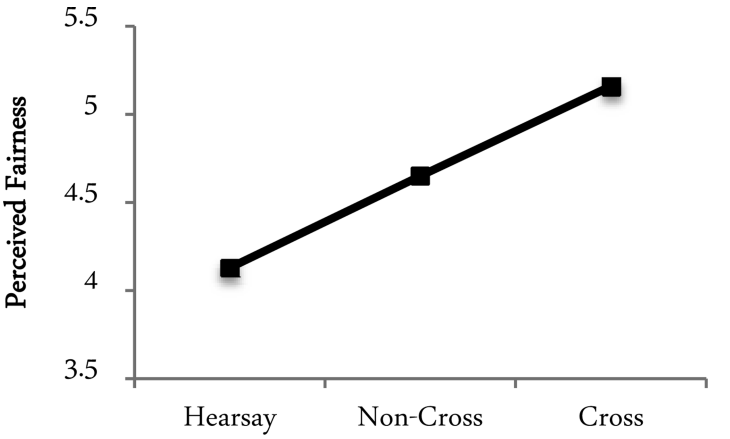
129. $F(1, 318) = 9.90$, $p = .002$.

130. To examine this question, I used a statistical technique referred to as a logistic regression analysis. A multiple regression analysis is a statistical test that estimates the independent effects of several predictor variables on a continuous dependent variable. A logistic regression is a regression analysis that examines whether several variables independently predict a binary, dichotomous outcome, such as being satisfied or unsatisfied. *See* LAWLESS ET AL., *supra* note 125, at 343–50.

131. $M_{\text{criminal}} = 4.13$ ($SD = 1.88$), $M_{\text{no-cross}} = 4.55$ ($SD = 1.74$), $M_{\text{cross-examination}} = 5.16$ ($SD = 1.88$); $F(2, 314) = 10.38$, $p < .001$, $\eta^2_p = .06$.

132. $F(2, 317) = 16.50$, $p < .001$.

Figure 5: Perceptions of the Fairness of the Procedure Employed in the Dispute



3. Accuracy

If participants' antipathy toward hearsay evidence is primarily a function of the unfairness of presenting hearsay evidence instead of its potential lack of accuracy, we would expect a different pattern of results with respect to participants' perceptions of the fact finder's ability to reach an accurate verdict. This hypothesis was tested first with an ANOVA, which revealed a statistically significant effect of the type of hearsay presented to the defendant.[133]

An ANOVA, however, reveals only whether any of the means in the experimental conditions differ from each other. A contrast analysis determines whether the means form a linear pattern. Unlike participants' satisfaction ratings and fairness judgments, a linear contrast analysis of participants' perceptions of decisional accuracy revealed no linear pattern in Figure 6,[134] and a further analysis indicated that the means in the hearsay confession and live witness (no cross-examination) conditions were not different from each other, whereas the live witness (with cross-examination) condition produced greater perceptions of the decisional accuracy of the administrative hearing and the criminal tribunal.[135] In sum, producing a witness did not alter participants' perceptions of decisional accuracy; therefore a belief that the mere presence of a witness might allow for a fact finder to evaluate that witness' demeanor is unsupported by the data.[136]

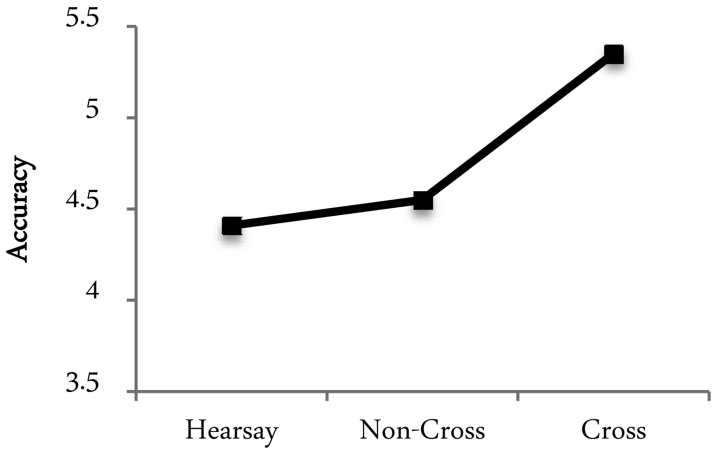
133. $M_{\text{hearsay}} = 4.41$ ($SD = 1.81$), $M_{\text{no-cross}} = 4.55$ ($SD = 1.67$), $M_{\text{cross-examination}} = 5.35$ ($SD = 1.89$); $F(1, 315) = 9.87$, $p < .001$, $\eta^2_p = .06$.

134. $F(1, 318) = 2.49$, $p = .115$.

135. For the comparison of the hearsay condition to the no-cross-examination condition, $p = .576$; for the comparison of the no-cross-examination condition to the cross-examination condition, $p = .001$.

136. A mediation analysis, performed below in Part III.C.4, confirms this result.

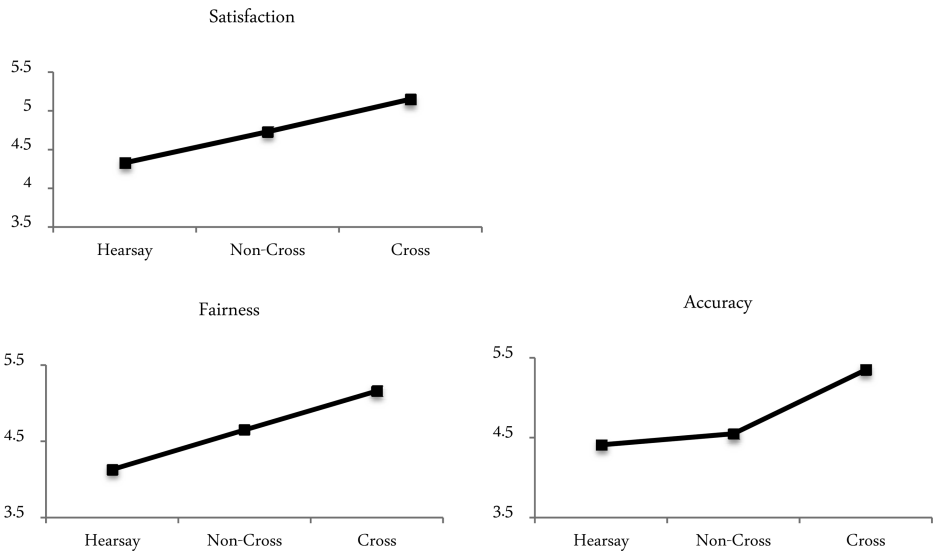
Figure 6: Perceptions of the Accuracy of the Procedure Employed in the Dispute



4. Mediation Analysis

The experimental results so far suggest that participants' satisfaction with the evidence presented to them followed a linear pattern. They were least satisfied with a hearsay confession, more satisfied with the presence of a live witness, and most satisfied with the presence of a live witness who was subject to cross-examination.

Figure 7: Perceptions of General Satisfaction, Fairness, and Accuracy by Condition



As illustrated in Figure 7, the results so far also shed light on the reason why participants prefer trials with a live witness subject to cross-examination over trials that rely on a hearsay confession. Recall that two possibilities emerge. The first is that tribunals in which live witnesses testify will reach more accurate decisions because the evidence has been vetted in the adversary process. The second is that the live witness provides the defendant with an important dignity interest—getting to "look his accuser in the eye"—apart from decisional accuracy.

The data so far show the participants' satisfaction ratings and fairness ratings exhibited a linear pattern. However, the participant's ratings of the accuracy of the dispute resolution procedure to which they were exposed did not exhibit a linear pattern. This suggests that participants' satisfaction with the proceeding to which they were exposed was more likely to be a function of the perceived fairness of the proceeding and not its ability to make accurate decisions. Psychologists can test this hypothesis empirically through a procedure called a mediation analysis.

A mediation analysis is, at its core, a series of predictive "connect-the-dots" statistical statements.[137] A statistically meaningful mediation analysis first demonstrates that the predictor variable predicts the outcome variable.[138] It then demonstrates that the predictor variable also predicts a mediator variable, and that the mediator, in turn, predicts the outcome variable.[139] Finally, the mediation analysis demonstrates that the effect of the predictor variable on the outcome variable is reduced or eliminated when the mediator is added to the statistical model, which suggests that the mediator is responsible for the effect of the predictor variable on the outcome variable.[140]

Under this framework, I used a mediation analysis to determine whether the fairness of the proceeding, the accuracy of the decision, or both mediated the effect of the type of evidence to which participants were exposed on their satisfaction for the proceeding in which that evidence was produced. Initially, I confirmed the results of Study 1 by examining via a regression analysis whether the type of evidence to which participants were exposed (specifically, a hearsay confession, a live witness who was not cross-examined, or a live witness who withstood cross-examination) affected the participants' satisfaction with the

137. Justin Sevier, *Redesigning the Science Court*, 73 MD. L. REV. 770, 824 (2014). Mediation analysis detects "when a predictor affects a dependent variable indirectly through at least one intervening variable, or mediator." Kristopher J. Preacher & Andrew F. Hayes, *Asymptotic and Resampling Strategies for Assessing and Comparing Indirect Effects in Multiple Mediator Models*, 40 BEHAV. RES. METHODS 879, 879 (2008). The mediation analysis reported in this Article is performed using a linear regression analysis and reports unstandardized coefficients, "B," and standard errors, "SE." It also reports a "t" statistic, which determines whether the coefficients are statistically significant. A linear regression is a statistical test that estimates the independent effects of several predictor variables on a continuous dependent variable. *See* LAWLESS ET AL., *supra* note 125, at 29, 300–31.

138. Preacher & Hayes, *supra* note 137, at 879; Sevier, *supra* note 137, at 824.

139. Preacher & Hayes, *supra* note 137, at 879; Sevier, *supra* note 137, at 824.

140. Preacher & Hayes, *supra* note 137, at 879; Sevier, *supra* note 137, at 824.

proceedings. The analysis confirmed that a strong relationship exists between these variables.[141]

Next, I examined the role of perceived fairness to the defendant in mediating the effect of the type of evidence on participants' satisfaction with the proceeding. First, I evaluated whether the evidence to which participants were exposed affected their perceptions of fairness, and the result was statistically significant.[142] As the evidence became less like hearsay and more like a live cross-examination, perceptions of perceived fairness increased. Second, I examined whether the perceived fairness of the proceeding affected participants' judgments of satisfaction. The results indicated that, to a statistically significant degree, as perceptions of the fairness of the proceeding increased, judgments of satisfaction increased as well.[143] Altogether, the analysis confirms that the evidence to which participants were exposed predicted their perceptions of the fairness of the proceeding, and these perceptions of fairness, in turn, predicted their satisfaction with the proceeding.

I then examined the role of the ability of the court to produce an accurate verdict in mediating the effect of the type of evidence on participants' satisfaction with the proceeding. As with perceived fairness, I performed a regression analysis to determine whether the evidence to which participants were presented affected their perceptions of the tribunal's ability to render an accurate verdict. This time, and unlike the regression on participants' perceptions of fairness, this regression revealed no relationship between these two variables.[144]

As illustrated in the figure above, whereas the perceived fairness of the proceeding was a statistically significant mediator of the effect of the evidence on participants' satisfaction with the proceeding—in fact, it accounts for all of the effect of hearsay evidence on participants' satisfaction judgments—the perceived ability of the tribunal to render an accurate verdict given the evidence was not a statistically significant mediator.[145]

---

141. $B = .408$, $SE = 0.13$, $t = 3.15$, $p = .002$.

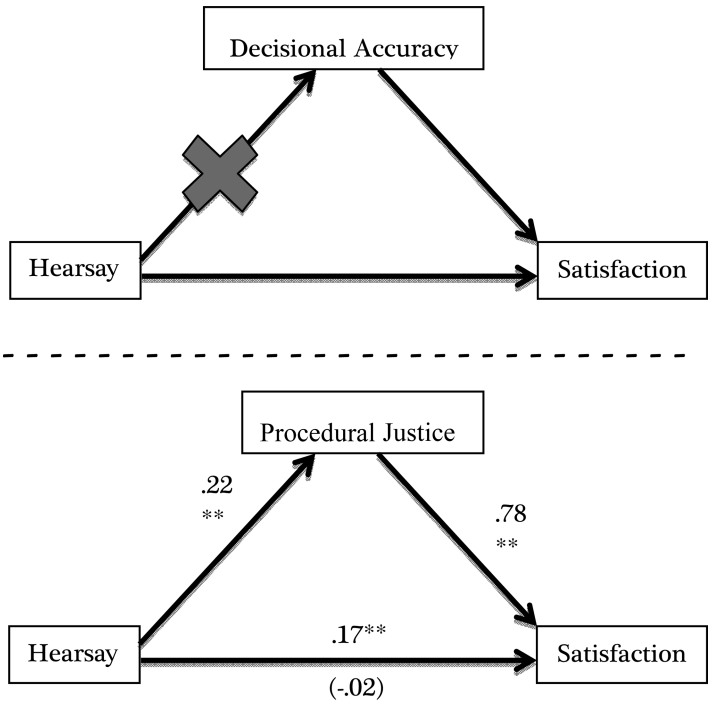
142. $B = .516$, $SE = 0.13$, $t = 4.06$, $p < .001$.

143. $B = .787$, $SE = 0.04$, $t = 22.13$, $p < .001$.

144. $B = .473$, $SE = 0.13$, $t = 3.82$, $p > .05$.

145. The effect of the evidence to which participants were exposed on participants' judgments of satisfaction, without the perceived fairness of the evidence to the defendant, is .17, which represents a standardized beta coefficient. When the perceived fairness of the evidence to the defendant is included in the model, the effect of the type of evidence is reduced to nonsignificance (a -.02 beta coefficient with a corresponding p-value above .05). Thus, the perceived fairness of the evidence to the defendant is a full mediator of participants' satisfaction with the proceeding. In contrast, with participants' perception of the ability of the tribunal to reach an accurate decision, the evidence was not a statistically significant mediator of the effect of the evidence on participants' satisfaction judgments.

Figure 8: Mediation Analysis Examining Which Variables Mediate the Effect of the Trial Procedure on Participants' Feelings of Satisfaction



The same results are evident when satisfaction is measured dichotomously (that is, when participants answer that they were "generally satisfied" or "generally dissatisfied" with the dispute resolution proceeding). The table below illustrates that, just like when satisfaction is measured on a seven-point scale, participants' perceptions of the evidence's fairness to the defendant—and not their perceptions of the ability of the court to make an accurate decision given the evidence that was presented—fully mediated participants' satisfaction with the proceeding.[146]

D. DISCUSSION

The experimental results reported in Study 1 reveal several findings regarding the public's perceptions of hearsay evidence at trial. First, the degree to which

146. A logistic regression is a regression analysis that examines whether several variables independently predict a binary, dichotomous outcome, such as a guilty or not guilty verdict. *See* LAWLESS ET AL., *supra* note 125, at 345–50 (discussing "logit regressions"). The odds ratio on the right side of the table represents the odds that an individual would choose one of the binary options given the effect of the independent variable on the left side of the graph. Thus, procedural justice fully mediated the effect of the hearsay evidence on participants' judgments of satisfaction, and for each one-unit increase in their perceptions of the procedural justice of the trial, they became 36% more likely to say that they were satisfied.

Table 1: Effect of Evidence, Procedural Justice, and Decisional Accuracy on Satisfaction

	B	SE	Wald	p-value	Odds
Hearsay Condition	—	—	0.43	.808	—
No Cross	0.03	0.29	0.01	.933	1.03
Cross-Examination	−0.15	0.28	0.26	.610	0.87
Procedural Justice	**0.31**	**0.11**	**8.55**	**.003**	**1.36**
Decisional Accuracy	−0.08	0.11	0.53	.465	0.93
Constant	−0.91	0.41	4.99	.025	0.41
Model χ^2	17.29*				
Pseudo R^2	.170				

Note: Responses were coded as 0 for unsatisfied. * indicates $p < .001$ with respect to model fit.

participants are concerned about hearsay evidence depends, partly, on the stakes of the dispute resolution procedure. Participants did not find hearsay to be overly problematic generally. Rather, participants became more troubled by hearsay evidence as the crime for which the defendant was accused—and his potential punishment—increased. Thus, although participants in this study appeared troubled by hearsay in both types of dispute resolution proceedings, they were more troubled when the hearsay evidence was used to accuse the defendant of conspiracy to commit murder than when the defendant was accused of a lesser type of conspiracy. There could, of course, be other reasons for this difference. For example, a courtroom jury may simply respond to hearsay evidence differently compared to a fact finder in an administrative hearing. This finding should therefore receive additional empirical attention.

More importantly, Study 1 sheds light on the reason for which participants expressed dissatisfaction with trials that involved hearsay evidence. The results suggest that, in both the criminal trial and the internal dispute proceeding, participants were least satisfied with a proceeding that relied primarily on hearsay evidence, more satisfied with a proceeding in which the witness appeared live without cross-examination, and most satisfied with a proceeding in which the live witness was cross-examined. If we maintain the belief that hearsay evidence is problematic only because of its effect on the decisional accuracy of a dispute resolution tribunal—because unreliable hearsay evidence might be afforded greater probative weight than it deserves—the results reported in Study 1 would support the view that participants are acting irrationally. There would be little difference, from the perspective of decisional accuracy, between a witness whose out-of-court statement is read into the record and a witness who appears live to read that same statement without being

cross-examined.[147]

A closer evaluation of the data suggests that participants were acting rationally. If participants' discomfort with the hearsay evidence to which they were exposed stems from a belief that it is unfair for out-of-court statements to be used against the defendant—that is, that the accuser should be required to appear in court to look the defendant in the eye—then it is rational that their discomfort with hearsay evidence would decrease when the accuser appears live at trial, even if the accuser provides the same information that would appear in the out-of-court statement.

Indeed, the mediation analysis makes clear that that the underlying rationale for participants' discomfort with hearsay evidence stems from their view that such evidence is unfair to the defendant, not that the evidence raises the risk of inaccurate verdicts. Participants' ratings of the fairness of the proceeding followed the same linear pattern as their satisfaction with the proceeding. But their view of the risk of an inaccurate verdict did not. Rather, participants' perception of the risk of an inaccurate verdict only decreased when the live witness was subjected to cross-examination; it was static in the remaining two experimental conditions. Thus, it is unsurprising that the mediator of participants' dissatisfaction with hearsay evidence was not decisional accuracy, but the dignity interest of the defendant on trial.

IV. STUDY 2

Study 1 suggests that laypeople are generally averse to hearsay evidence. Laypeople instead appear to be more satisfied with dispute resolution procedures that include live witnesses who withstand adversarial cross-examination. Study 1 also strongly suggests that laypeople do not see the legal prohibition on hearsay evidence as a rule that primarily safeguards the tribunal's decisional accuracy and minimizes the risk of erroneous verdicts. Rather, Study 1 suggests that laypeople perceive the rule against hearsay evidence as a rule that safeguards the dignity interests of the defendant and preserves fair play during the dispute resolution proceeding.

Participants in Study 1, however, were not directly asked for the reasons that underlie their discomfort with hearsay evidence. Although the mediation analysis strongly suggests that procedural fairness—and not decisional accuracy—is the primary reason for their discomfort, psychologists have methods by which they are able to measure participants' free responses to this question.

147. It would be an overstatement, however, to conclude that decisional accuracy is irrelevant to participants' satisfaction with hearsay. Although it does not explain the pattern of participants' satisfaction levels to a significant degree generally—and not as well as participants' perceptions of procedural justice—participants' levels of satisfaction and perceptions of decisional accuracy did increase in the cross-examination condition (compared to the condition in which the witness was live but was not cross-examined).

Study 2 therefore seeks to replicate the results found in Study 1 while also adding to those results in several ways. In Study 2, participants were exposed to a summary of the trial of Sir Walter Raleigh.[148] This time, there were two versions of the trial: (1) a version in which the alleged coconspirator's evidence against Raleigh took the form of an out-of-court hearsay confession (which is what occurred at the trial); and (2) a version in which Raleigh's coconspirator testified live to the same evidence and was not exposed to cross-examination. Thus, the same information was presented to participants and only the form of that information varied between the two conditions.

In Study 2, however, instead of being exposed to just one version of the trial, participants were exposed to both versions of the trial, in random order. In this experimental design, participants were able to compare the two versions of the Raleigh trial, decide what satisfied or dissatisfied them about those versions of the trial, and then discuss their satisfaction or dissatisfaction in their own words to the experimenters.

If participants dislike hearsay evidence, as the results of Study 1 suggest, then we would expect that participants would be significantly more satisfied with the version of the Raleigh trial in which the coconspirator testified compared to the version in which the hearsay statement was admitted into evidence. Moreover, we should be able to determine, in the participant's own words, whether the increase in perceived satisfaction is the result of perceptions of fairness or perceptions of increased decisional accuracy.

A. PARTICIPANTS

One hundred sixty-four participants were recruited for an online study using the web-recruiting service MTurk, which was also used in Study 1.[149] The sample of participants in this study was 53.00% female, 68.25% Caucasian, and averaged 37.18 years of age (with a standard deviation of 14.70 years). Seventy-one percent of the sample had completed at least a college degree, and the average household income of the sample was between $40,000 and $49,999.

B. PROCEDURE AND MEASURES

In Study 2, participants read two variations of a summary of the trial of Sir Walter Raleigh.[150] After reading one version of the trial, participants answered

148. For a discussion of the trial of Sir Walter Raleigh, see *supra* notes 63–66 and accompanying text.

149. *See supra* note 122.

150. In Study 1, participants were randomly assigned to just one experimental condition, and the responses from different participants, who had been assigned to different experimental conditions, were compared in a "between subjects" design. In contrast, in Study 2, each participant was exposed to both experimental conditions and each participant's responses to each condition were measured in a "repeated measures" design. A repeated measures design has many benefits for statistical hypothesis testing, which includes reducing the global variability and error in the statistical analysis, because each participant provides data for two different experimental conditions. *See, e.g.*, DAVID C. HOWELL, STATISTICAL METHODS FOR PSYCHOLOGY 462 (7th ed. 2010).

several questions. First, they answered a series of questions designed to gauge the participants' attention to and recollection of the trial vignette. After successfully answering those questions, they answered several questions designed to gauge their satisfaction with the case, on the same seven-point scale that was used in Study 1. Additionally, participants were asked whether they were generally satisfied or dissatisfied with the trial.

Next, participants rated, on a seven-point scale, the degree to which they believed that the trial was fair and the degree to which the court was able to accurately uncover the facts of the case. Unlike in Study 1, however, participants did not answer similar questions about the individual pieces of evidence, because the trial in Study 2 consisted solely of the statements of Raleigh's alleged coconspirator.[151] Afterward, participants read the second variation on the Sir Walter Raleigh trial and answered the same questions.

When participants read the trial variation that involved the hearsay confession—which was the first version for half the participants and the second version for the other half—they were asked additional questions. If participants responded that they were dissatisfied with the trial that contained the hearsay confession, they were asked to type (in as many words as they preferred) the reason for their dissatisfaction.[152]

After stating the reason for their dissatisfaction in their own words, participants were shown a list of potential rationales for being dissatisfied with the trial. These rationales included: (1) that we cannot determine the accuracy of the confession; (2) that it is unfair to the defendant for an accuser's words to be used against him without showing up in court; and (3) none of the above. Participants were asked to choose the option that best encapsulated the reason for their dissatisfaction with the trial that contained hearsay evidence. If participants chose the "none of the above" option, they were asked to type in the reason for their dissatisfaction.

Next, participants were asked to rate, on a seven-point scale anchored at "it is mostly about factual accuracy" on the left and "it is mostly about fairness" on the right, the reason for which a judge would exclude the evidence. This question was designed to measure the strength of each participant's preference for an "accuracy" or "fairness" rationale.

Finally, participants were asked to provide demographic information, including their age, race, income, level of education completed, and their familiarity

---

The order in which participants read about the two trials was randomized. This is standard protocol in a repeated measures design, because experimenters would otherwise raise the risk of order effects, wherein participants' data are systematically affected by whether they were exposed to a particular experimental condition first or last. Randomizing the order in which participants were exposed to the different versions of the trial eliminates potential order effects. *See, e.g.*, Robin M. Hogarth & Hillel J. Einhorn, *Order Effects in Belief Updating: The Belief-Adjustment Model*, 24 COGNITIVE PSYCHOL. 1 (1992).

151. In other words, to the extent that participants' general satisfaction with the trial and their perceptions of its fairness and accuracy are evidence-based, these could be affected only by the information provided by the alleged coconspirator.

152. Participants' responses were then coded in a procedure discussed more fully below.

with the legal system. Participants were then debriefed with respect to the aims of the study, and the experiment was concluded.

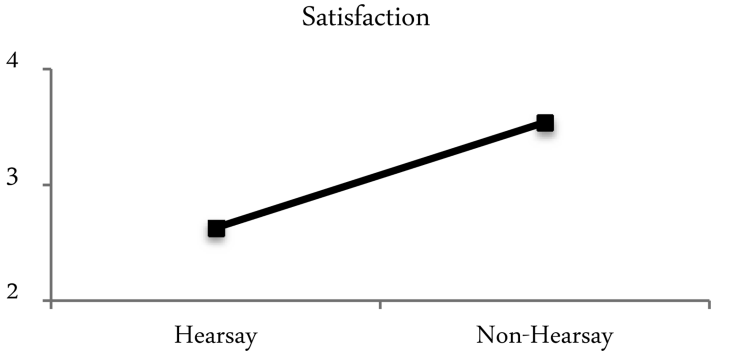
C. RESULTS

The analysis of the results in Study 2 proceeds in two parts. First, it attempts to replicate the results found in Study 1 that (1) judgments of satisfaction will be lower when hearsay is presented than when it is not and (2) those satisfaction judgments are attributable to the perceived unfairness of the evidence to the defendant and not to any perceived decisional accuracy concerns. Second, it seeks to provide additional evidence, in the form of free-response data and forced-choice data, that the concerns participants have with hearsay evidence are the result of hearsay's perceived unfairness to the defendant.

1. Two Trials

First, I examined whether statistically significant differences in perceived satisfaction existed when participants read the summary of the Sir Walter Raleigh trial and the coconspirator's evidence took the form of an out-of-court confession, or that same evidence was presented in the form of the coconspirator testifying live at trial (but not subject to cross-examination). An ANOVA revealed a significant effect of the form of the evidence, such that participants reported higher satisfaction with the trial when the witness testified live than when the witness's testimony took the form of an out-of-court hearsay confession.[153]

Figure 9: Ratings of Self-Reported Satisfaction by Experimental Condition

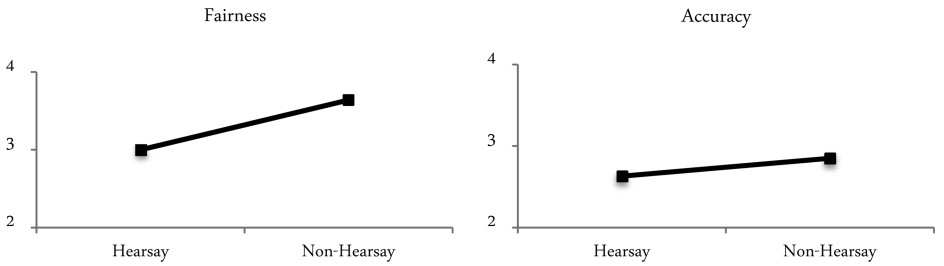


153. $M_{\text{hearsay}} = 2.63$ ($SD = 1.69$), $M_{\text{non-hearsay}} = 3.54$ ($SD = 1.85$); $F(1, 159) = 10.77$, $p = .001$, $\eta^2_p = .06$. The same result occurred when the outcome variable was a dichotomous measure of satisfaction (that is, satisfied or unsatisfied). A logistic regression revealed that participants were two times more likely to be satisfied with the proceedings when the evidence was presented live than when it was presented in the form of a hearsay statement. $B = .699$, $SE = 0.40$, $Wald = 3.01$, $p = .082$ (marginal).

Next, I examined whether statistically significant differences in participants' perceptions of the fairness to the defendant existed when the testimony against him was presented live instead of as a hearsay statement. An ANOVA again revealed a statistically significant effect of the form of the evidence, such that participants perceived the trial to be fairer to the defendant when the coconspirator testified live in court than when the coconspirator's information was conveyed in a hearsay statement.[154]

I then performed this same analysis with respect to participants' perceptions of the tribunal's ability to reach an accurate decision. This time, however, the ANOVA revealed no effect of the form of the evidence. Participants perceived no difference in the ability of the court to reach an accurate verdict when the witness's testimony was presented live (without cross-examination) instead of when it was presented as an out-of-court hearsay statement.[155] A side-by-side comparison of the results presented thus far is presented in Figure 10.

Figure 10: Ratings of the Fairness and Accuracy of the Proceedings by Experimental Condition



Next, as in Study 1, I conducted a mediation analysis to determine whether the differences in participants' perceptions of the fairness of the trial affected their satisfaction with it. Recall that Figures 9 and 10 revealed that the form of the evidence predicted participants' satisfaction with the evidence and the participants' perceptions of the fairness of the trial to the defendant. A mediation analysis further reveals that participants' perceptions of the fairness of the trial predicted their satisfaction with the trial, such that perceptions of greater fairness to the defendant were associated with higher satisfaction with the trial.[156]

In contrast, the results reported in Figure 9 have already established that perceptions of the tribunal's ability to produce an accurate verdict were not

154. $M_{\text{hearsay}} = 3.00$ ($SD = 1.53$), $M_{\text{non-hearsay}} = 3.54$ ($SD = 1.85$); $F(1, 159) = 4.10$, $p = .044$, $\eta^2_p = .03$.

155. $M_{\text{hearsay}} = 2.63$ ($SD = 1.69$), $M_{\text{non-hearsay}} = 2.85$ ($SD = 1.70$); $F(1, 159) = 1.42$, $p = .236$, $\eta^2_p = .01$.

156. $B = .958$, $SE = 0.04$, $t = 26.04$, $p < .001$.

associated with the form of the evidence, although higher perceptions of accuracy were generally associated with greater satisfaction.[157] Thus, participants' perception of the tribunal's ability to produce accurate verdicts is eliminated as a potential mediator for the effect of the form of the evidence on participants' satisfaction with the trial.

Finally, an evaluation of a regression model that includes the form of the evidence and perceptions of fairness to the defendant as predictor variables, and the seven-point scale representing participants' satisfaction with the trial as the dependent variable, shows that perceptions of fairness significantly mediated the effect of the form of the evidence on participants' satisfaction with the trial.[158]

2. In Their Own Words

As reported in the previous section, the proportion of satisfied and dissatisfied participants was significantly different across the different versions of the trial. Specifically, although participants were generally more dissatisfied than satisfied with both versions of the trial, a meaningfully greater proportion of participants were dissatisfied in the hearsay version of the trial than in the nonhearsay version.[159] Figure 11 below illustrates the percentage of participants

157. $B = .822$, $SE = 0.05$, $t = 15.78$, $p < .001$.

158. When the perceived fairness of the evidence to the defendant is included in the model, the effect of the type of evidence is reduced from .252 to .118, which is a 53% reduction. Thus, the perceived fairness of the evidence to the defendant accounted for 53% of the variability in participants' judgments of satisfaction with the proceeding. In contrast, participants' perception of the ability of the tribunal to reach an accurate decision given the evidence was not a statistically significant mediator of the effect of the evidence on participants' satisfaction judgments. Moreover, when the outcome variable is, instead, participants' dichotomous satisfaction choice, the same result is revealed. First, participants' dichotomous satisfaction choice is highly correlated with their scaled satisfaction scores, $r(160) = .69$, $p < .001$. More importantly, when perceptions of procedural justice are added to the model as a mediator, the effect of the hearsay condition is reduced to nonsignificance, and increases in the perceived procedural justice made participants four times more likely to be dissatisfied with the trial that contained hearsay evidence. Decisional accuracy, in contrast, did not mediate the effect. The logistic regression table is presented below:

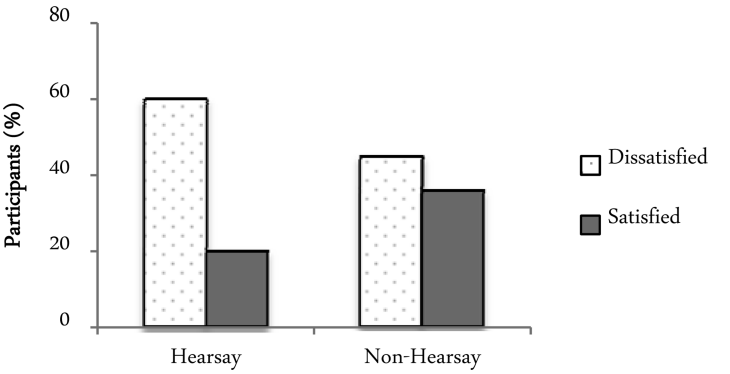
	B	**SE**	**Wald**	**p-value**	**Odds**
Hearsay	−0.96	0.67	2.04	.153	0.38
Procedural Justice	**1.65**	**0.37**	**19.94**	**<.001**	**5.19**
Decisional Accuracy	0.36	0.33	1.17	.279	1.43
Constant	−8.19	1.70	23.29	<.001	0.00
Model χ^2	98.01*				
Pseudo R^2	.722				

Note: Responses were coded as 0 for unsatisfied. * indicates $p < .001$ with respect to model fit.

159. $B = .699$, $SE = 0.40$, $Wald = 3.01$, $p = .082$ (marginal).

who were dissatisfied or satisfied by both versions of the trial of Sir Walter Raleigh:

Figure 11: Percentage of Satisfied and Dissatisfied Participants by Condition



To the extent that participants were unhappy with the version of the Raleigh trial that contained hearsay evidence—which reflects the participants in the far left bar in the figure above—Study 2 asked them specific questions designed to understand the reason for their unhappiness. Thus, Study 2 attempted to elicit, in the participants' own words, the mediator (or mediators) of their unhappiness with the Sir Walter Raleigh trial.

Participants' responses to the free-response items were evaluated through a procedure called *a priori* theme coding. Theme coding involves creating distinct categories for each participant's response to questions and determining the distinct words or phrases that, holistically, would cause a participant's response to fall into one distinct category.[160] The coding scheme for each particular category is then created, and two coders, who do not know the experimental hypothesis, evaluate each participant's response and assign it to one of the categories.[161]

Study 2 utilized four categories of responses: (1) a response that implicates decisional accuracy; (2) a response that implicates fairness to the defendant; (3) a response that implicates both decisional accuracy and fairness; and (4) a response that implicates neither of these phenomena.

160. For a thorough review of content analysis, see Steve Stemler, *An Overview of Content Analysis*, 7 PRAC. ASSESSMENT, RES. & EVALUATION 17 (2001), http://pareonline.net/getvn.asp?v=7&n=17.

161. If the responses require it, the experimenters can adjust the coding scheme to include additional categories, and the process occurs again. *See id*. Theme coding invariably involves a degree of subjectivity, insofar as coders make holistic judgments about how to interpret a participant's written response. *See id*. The subjectivity is minimized, however, by requiring independent coders to code all of the data separately. To the extent that the independent coders agree on how to classify each participant's response, this serves as convergent evidence that the classification is justified.

Responses were coded as implicating decisional accuracy if they used certain dictionary terms that are related to accuracy or truth.[162] They were coded as implicating procedural justice if they used dictionary terms related to fairness or justice.[163] Responses that listed both decisional accuracy and procedural justice (or equally) were marked as implicating both truth and justice, and responses that did not list any of these terms were marked as implicating neither. These responses were then re-evaluated to determine whether they implicated another relevant psychological construct. A sample of participants' responses—and how they were coded—appears in the table below:

Table 2: Participants' Explanation for Dissatisfaction with the Trial

Sample "Accuracy" Responses	Sample "Justice" Responses
"If we can't see the guy, how do we know if he's telling the truth?"	"It's unfair for the defendant not to be able to see the witness in court."
"The jury might make the wrong decision unless we see the man who wrote the confession."	"It just seems wrong to me that someone can be convicted without the person accusing them showing up in court."
"If we want to figure out what actually happened, the person who conspired with him has to show up, too."	"I think the court is violating the rights of the defendant by not forcing the other person to appear in court."
"Cross-examination is necessary to see if the other guy's story makes sense or not."	"We have fair trials in this country, which means we don't let people make unfounded accusations against us."
"You have to hear from everyone involved if you want to make the right decision in the case, I'd think."	"The Constitution requires prosecutors to give defendants a chance to see the people being used to put them away."

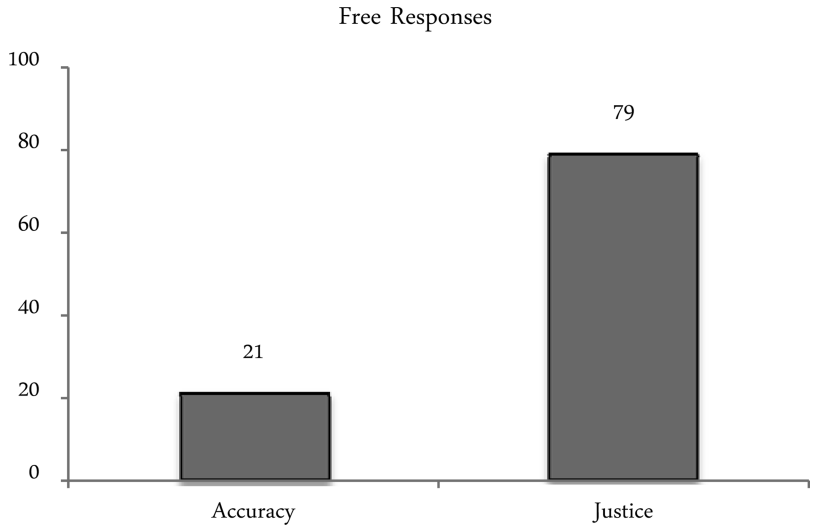
Moreover, as the figure below illustrates, 21% of the responses from partici-

162. These terms were based off a scale used in another work by the author and included "fact," "certainty," "reality," "actuality," "veracity," "verity," "accuracy," "honesty," "genuineness," "precision," or "exactness." *See* Justin Sevier, *The Truth-Justice Tradeoff: Perceptions of Decisional Accuracy and Procedural Justice in Adversarial and Inquisitorial Legal Systems*, 20 PSYCHOL. PUB. POL'Y & L. 212, 214 (2014). Responses were also coded as implicating decisional accuracy if they included phrases such as "accurate decisions," "correct decisions," "correct answers," "to get at the truth," "to find out what really happened," "to uncover the truth," "to determine who is right and who is wrong," "to get at the facts," "to reveal the right information," "to make a good factual decision," or "to resolve disputes correctly." *See id.*

163. These included terms such as "fairness," "impartiality," "righteousness," "reasonableness," "evenhandedness," "integrity," "uprightness," "rightness," or "just." Responses were also coded as implicating fairness to the defendant if they included phrases such as "to provide justice," "to give people a fair trial," "to give people control over their fate," "to allow people to present their arguments," "to allow people to make their points to a neutral (or unbiased) party," "to protect people's rights," "to give people dignity or respect," "to get a fair result," "to create a compromise that the parties can live with," or "to punish people appropriately." *See id.*

pants who were unhappy with the trial of Sir Walter Raleigh could be categorized as implicating either the decisional inaccuracy associated with hearsay evidence or unfairness to the defendant. But consistent with the statistical results reported in Study 1 and Study 2, participants' free responses indicate that 79% of participants stated their unhappiness in terms of the fairness to the defendant.[164]

Figure 12: Percentage of Dissatisfied Participants Discussing Accuracy or Justice Concerns



D. DISCUSSION

Study 2 yields several findings that enhance our understanding of how the public evaluates hearsay evidence and the reasons that underlie these evaluations. First, Study 2 replicates the results reported in Study 1. When participants were exposed to two different versions of the trial of Sir Walter Raleigh, they were significantly more satisfied with the version of the trial in which Raleigh's coconspirator testified live (without being cross-examined) compared to the version in which the coconspirator's evidence took the form of an out-of-court

164. The free response items converge with other measures used in Study 2 to assess the phenomena that mediate participants' disapproval of hearsay evidence. When asked to choose between a decisional accuracy rationale and a procedural justice rationale for their dissatisfaction with the trial that involved hearsay, the majority of participants chose the fairness rationale (with 70% choosing procedural justice and 30% choosing decisional accuracy).

Moreover, when participants had to choose, on a six-point scale with no midpoint, which rationale most closely expressed their dislike for the hearsay evidence (anchored at 1 = decisional accuracy and 8 = procedural justice), the average response was fairness-oriented and not accuracy-oriented (M = 4.84, SD = 1.75).

hearsay statement. This is so even though exactly the same information was conveyed in both trials, and the probative value of the evidence did not change. As in Study 1, if we focus solely on hearsay as a doctrine that promotes decisional accuracy, participants' greater satisfaction with a trial in which there is no increase in the probative value of the information provided to the tribunal appears puzzling. However, Study 2 also reveals that participants' perception of the court's ability to produce an accurate decision was statistically the same in the hearsay version of the trial as in the nonhearsay version, which suggests that participants understand that the probative value of the evidence is unchanged.

Thus, it does not appear to be true that participants prefer live witnesses to ensure decisional accuracy. Instead, as in Study 1, the results from Study 2 suggest that the increase in satisfaction that participants felt when they were exposed to the version of the trial in which the witness testified was attributable to their perception that this type of trial is more fair to the defendant than a trial in which an out-of-court witness accuses the defendant without testifying in person.

Moreover, participants in Study 2 explicitly confirmed this proposition in their own words. When participants were asked why they were dissatisfied with the Raleigh trial that contained hearsay evidence, the vast majority of participants stated that the reason for their discomfort with hearsay evidence involves procedural justice concerns.[165] In sum, these participants believed that accusatory witnesses should testify in-person at trial, not because doing so would reduce the risk of inaccurate verdicts, but because the defendant has a dignity interest in being convicted by the testimony of a witness who will look him in the proverbial eye.

V. IMPLICATIONS, OBJECTIONS, AND CONCLUSIONS

The most important, and unanswered, empirical query with respect to hearsay evidence is to determine what values the public believes the hearsay rule is designed to protect. The primary rationale for the rule barring hearsay evidence focuses on the extent to which barring potentially unreliable evidence decreases the risk of inaccurate judicial fact-finding. The difficulty with this rationale is two-fold, however. First, the empirical research conducted to date suggests that the detrimental effects of hearsay evidence on the decisional accuracy of a decision-making body might be a figment of rule makers' collective imagination. Instead, the social-science evidence suggests that jurors are far more likely to discount hearsay compared to other types of evidence because they appar-

165. Most participant responses were classified as implicating accuracy or justice concerns by the independent coders. Any disagreements were resolved through discussion, and the results reported here reflect postdiscussion classifications. These results align with empirical work that suggests that, although laypeople do perceive a relationship between accuracy and justice—insofar as one may incorporate the other—they also perceive them as distinct concepts. *See* Tom R. Tyler & Justin Sevier, *How Do the Courts Create Popular Legitimacy?: The Role of Establishing the Truth, Punishing Justly, and/or Acting Through Just Procedures*, 77 ALB. L. REV. 1095, 1095–98 (2014).

ently recognize that hearsay evidence is subject to several types of testimonial infirmities.[166]

A procedural justice rationale for the hearsay rule will have the support of the general public because dignity and fairness concerns—and not decisional accuracy concerns—are what the vast majority of the public believes the hearsay rule is designed to protect. Two original experiments reported in this Article have provided the first empirical support for the proposition that conceiving of the rule barring hearsay evidence as protecting the dignity concerns of litigants is likely to lead to greater popular legitimacy for the rule. The first study found that, in criminal and noncriminal contexts, participants were far less satisfied with the trial proceeding when hearsay evidence was present, and that their satisfaction was specifically tied, not to the ability of the tribunal to determine the truth of the matter given the evidence, but to their perception that the inclusion of that evidence was unfair to the defendant. This finding was confirmed in the second study, in which participants' candid responses of dissatisfaction with the trial of Sir Walter Raleigh overwhelmingly suggested that their dissatisfaction with the hearsay evidence was process-based. The vast majority of participants were dissatisfied with Raleigh's trial because the use of the hearsay confession to convict him seemed unseemly to participants; their dissatisfaction was not tied to any effects that the confession might have on the tribunal's ability to reach the correct substantive outcome.

Taken together, these studies provide support for the growing weight of empirical data challenging the appropriateness of relying on a decisional accuracy rationale to support the ban on hearsay evidence. The results also have implications for the future direction of the hearsay rule and for ground-level decisions that practicing attorneys make daily with respect to hearsay evidence.

A. RESEARCH AND POLICY IMPLICATIONS

It is becoming increasingly apparent that the decisional accuracy rationale for the hearsay rule is crumbling under the weight of empirical research. Empirical hearsay studies continue to converge on the same conclusion: jurors are significantly more competent to evaluate hearsay evidence than policymakers credit them to be.[167] In stark contrast, the procedural justice rationale for the rule

166. Moreover, a preoccupation with the potentially detrimental effects of hearsay on legal decision making can lead to empirical questions that are incapable of a satisfactory answer. To the extent that we fashion the hearsay doctrine according to a concern that jurors do not afford hearsay evidence exactly the probative weight that hearsay deserves raises the question of how to assign probative weight to evidence that is presented at trial at all. Because there exists no clear, objective standard by which judicial fact finders weigh the evidence that is presented to them at a legal proceeding, the proposition that hearsay evidence will not be discounted appropriately enough by judicial decision makers is unfalsifiable and counterproductive to creating coherent, evidence-based hearsay policy.

167. *See supra* Part II.C. It is, however, an open question—and one worth further empirical study—whether jurors are actually competent or whether they are simply more competent than we expect. As discussed earlier in this Article, a measurement problem exists that currently prevents us from further exploring which is the more accurate phrase.

barring hearsay evidence has significant support in both the theoretical literature and in empirical, social psychological studies.[168]

The implications for the hearsay doctrine are numerous. If jurors are competent to evaluate hearsay evidence, the decisional accuracy rationale for the hearsay rule must give way to a more persuasive justification. Some researchers have argued, based on the empirical data, that hearsay evidence should be allowed into court because hearsay often contains at least some informational value and jurors can be trusted to discount hearsay in a reasonable manner.[169] It is not obvious, however, that the solution is to eliminate the hearsay rule. Eliminating the hearsay rule altogether is likely to create greater cost to the legal system and to litigants than keeping the rule, but relying on a rationale that is theoretically coherent and empirically credible.

If we conceive of the hearsay rule as a rule that promotes fairness in the trial process—which the studies reported in this Article suggest is how the public conceives of the doctrine—we might argue that, as a matter of dignity, criminal and civil trial judges should not allow into evidence information from accusers whom the defendant has not had the opportunity to face and to cross-examine.[170] Hearsay policy that relies on principles that underlie the Sixth Amendment Confrontation Clause (and its state-law counterparts) shifts the normative debate from one that is almost entirely empirically based to one that is based both on empirical testing and sound psychological theory.

A procedural justice rationale would also more closely align with recent constitutional developments with respect to the hearsay doctrine. Although the controlling rationale for the U.S. Supreme Court's hearsay jurisprudence has become less predictable recently, its decisions in *Crawford v. Washington*[171] and *Davis v. Washington*[172] signal that the Supreme Court finds this procedural justice rationale appealing, although the Court also appears to be focused, in part, on empirical assertions about the power of cross-examination to expose testimonial infirmities.[173]

A shift to a procedural justice justification for barring hearsay evidence in civil and criminal cases would require policymakers to reevaluate the exceptions to the doctrine under Federal Rule of Evidence 803—which allows into evidence so-called reliable hearsay, even though the reliability of such evidence

168. *See supra* Part III.

169. *See* Swift, *supra* note 69, at 495–98; *see also* David Crump, *The Case for Selective Abolition of the Rules of Evidence*, 35 HOFSTRA L. REV. 585, 612–14 (2006); George F. James, *The Role of Hearsay in a Rational Scheme of Evidence*, 34 ILL. L. REV. 788, 790–94 (1940); Jack B. Weinstein, *Probative Force of Hearsay*, 46 IOWA L. REV. 331, 338–39 (1961).

170. *See supra* Part II.B.

171. 541 U.S. 36 (2004).

172. 547 U.S. 813 (2006).

173. *See Crawford*, 541 U.S. at 61 (discussing the testing of evidence through the "crucible of cross-examination"). It should be noted that, although the right of confrontation attaches in criminal trials as a matter of constitutional law, under a procedural justice rationale for the hearsay rule, a similar right could attach as a policy matter in civil trials as well.

currently has no empirical support—and under Federal Rule of Evidence 804—which allows into evidence potentially unreliable hearsay statements for fear of losing all evidence on a particular issue at trial.[174] These concerns would prove challenging, as would the concerns facing any coherent framework for justifying the bar on hearsay evidence. Any such rationale, however, would represent an improvement from the current state of the doctrine, in which hearsay is banned largely on account of folk wisdom about juror cognition that is unsupported by empirical research.[175]

The findings reported here also have implications for practitioners, who make ground-level decisions about hearsay evidence. There are myriad reasons for attorneys to use hearsay evidence instead of in-court testimony: a witness could no longer be alive or could be ill, she could have moved away from the jurisdiction, she might be unavailable for other important reasons, or she might refuse to testify.[176] The data presented here and elsewhere, however, suggest that jurors make strong judgments about an attorney's decision to use hearsay evidence—almost always negative—that may backfire on the attorney who calls the hearsay witness.[177] It therefore behooves attorneys to think critically about using hearsay testimony for reasons other than necessity. Any limitations in the declarant's testimony that the attorney may seek to obscure through hearsay might be ferretted out and weighed accordingly by jurors.

B. OBJECTIONS AND FUTURE DIRECTIONS

The findings reported in this Article provide new evidence that challenges the decisional accuracy rationale for the hearsay rule. Controlled behavioral experiments are, of course, subject to specific limitations of which policymakers should be mindful. Discussing these limitations—and potential responses to them—can increase policymakers' willingness to consider the data and can provide researchers with future avenues to explore with respect to the hearsay rule.

Controlled experiments differ from field studies in an important respect: field studies observe participants in their natural environments whereas laboratory experiments observe participants' behavior and judgment in a more uniform setting.[178] Neither design is superior to the other; they each involve important trade-offs that policymakers should consider. Field experiments have the benefit

174. *See supra* notes 105–11 and accompanying text.

175. *See supra* Part III.

176. *See* Michael L. Seigel, *Rationalizing Hearsay: A Proposal for a Best Evidence Hearsay Rule*, 72 B.U. L. REV. 893, 919 n.82 (1992); Sevier, *supra* note 107, at 2–3; *see also* Liza I. Karsai, *The "Horse-Stealer's" Trial Returns: How* Crawford*'s Testimonial-Nontestimonial Dichotomy Harms the Right to Confront Witnesses, the Presumption of Innocence, and the "Beyond a Reasonable Doubt" Standard*, 62 DRAKE L. REV. 129, 170 n.228 (2013).

177. *See, e.g.*, Sevier, *supra* note 107, at 1–2 (finding that jurors make motivational attributions for the inclusion or exclusion of hearsay evidence).

178. *See* Lynne ForsterLee & Irwin A. Horowitz, *The Effects of Jury-Aid Innovations on Juror Performance in Complex Civil Trials*, 86 JUDICATURE 184, 184–85 (2003); Jennifer K. Robbennolt,

of external and ecological validity.[179] Experimenters who use a field study methodology can assert with more confidence than an experimenter who employs a laboratory study that the results reflect what participants actually do in response to an environmental stimulus. There is always a concern in laboratory studies that the connection between what is found in the lab and what occurs in the real world might not be as strong as experimenters believe, although data collected on this issue suggest that this concern is overstated.[180] However, what field studies gain in external validity, they lose with respect to internal validity—the ability of the researcher to express with confidence that she measured in her study what she claims she measured.[181]

All else equal, controlled laboratory studies contain much greater internal validity than do field studies, because the environment in a laboratory study—whether it is a vignette design or a behavioral design—is kept uniform for all study participants with the exception of the experimental manipulation.[182] Thus, any differences observed among participants in an experiment are attributable solely to the experimental manipulation. This allows researchers in a controlled laboratory experiment to reach stronger causal conclusions—that manipulation X caused response Y—than researchers who employ other empirical methods. Although field researchers can (and do) attempt to control for potential confounding factors through the use of statistical techniques, many scholars agree that statistical controls are inferior to a randomly assigned experimental design in which the manipulation is all that differs among the experimental conditions.[183]

Moreover, true threats to external validity involve an interaction between the functional relationship being studied—that is, the effect of the independent variable on the dependent variable—and the setting—a laboratory simulation versus a field test. That the overall level of an effect is higher or lower in one setting compared to another is rarely a matter of concern among scientists; indeed, by definition it is not a concern when the question is whether a hypothesized functional relationship exists or not. External validity concerns arise when an independent variable increases a dependent variable in one setting but decreases it in the other setting. Were an effect found in one setting to merely disappear in the other setting, that might or might not be a concern from a policy perspective, but that would depend on the details of the policy question.

The vignettes employed in the studies reported here provide important information regarding how participants evaluate hearsay evidence. The trials in both

---

supra note 125, at 483; *see also* Justin Sevier, *The Unintended Consequences of Local Rules*, 21 CORNELL J.L. & PUB. POL'Y 291, 324–25 (2011).

179. Sevier, *supra* note 178, at 325 & n.249.

180. *See* Mark Kelman, Yuval Rottenstreich & Amos Tversky, *Context-Dependence in Legal Decision Making*, *in* BEHAVIORAL LAW AND ECONOMICS 61, 73 (Cass R. Sunstein ed., 2000).

181. *See supra* note 117 and accompanying text.

182. *See* LAWLESS ET AL., *supra* note 125, at 93–96.

183. *See* LAWLESS ET AL., *supra* note 125, at 93–122 (discussing the strengths, weaknesses, and tradeoffs among controlled laboratory experiments, field experiments, quasi-experiments, and natural experiments).

studies were the same in nearly every respect—except the hearsay manipulation—and revealed stark differences and significant trends among participants with respect to the manner in which they evaluate hearsay evidence. A field study design, in which different participants would likely be exposed to different cases with different facts, would not be able to produce statements of causality with respect to people's perceptions of hearsay evidence that are as strong as a laboratory design. Nonetheless, field studies of hearsay evidence—which have not yet been conducted by empirical researchers—should be the next step in gathering data about hearsay evidence. Using publicly available data, ambitious researchers could code real cases for the presence or absence of hearsay evidence, the type of hearsay that was submitted, and code for (and ultimately control for) factors such as the charges against the defendant, demographics of the relevant legal actors, and the complexity of the trial in order to draw conclusions from cases involving hearsay in the real world. The external and ecological validity reported by such cases would complement the internal validity supplied by controlled laboratory studies and, together, would supply convergent validity for how laypeople think about hearsay evidence.[184] At the least, other researchers should consider replicating the results reported in this Article in a videotaped trial or a live reenactment.

CONCLUSION

Evidence rule makers have created a hearsay policy quagmire in the absence of data that could assist them in determining whether the dangers they see in hearsay evidence have an empirical basis, and whether their understanding of the values that the hearsay doctrine is designed to protect match the values of the citizens they serve. The studies reported in this Article provide a framework for understanding how the public thinks about hearsay and challenge the common law rationale, conceptualized in Professor Tribe's testimonial triangle, that the hearsay rule promotes decisional accuracy.

Policymakers should consider the evidence reported in this Article when examining the concerns that the hearsay rule is designed to remedy. Did the outrage over the trial of Sir Nicholas Throckmorton stem from a belief that incompetent jurors could have clumsily evaluated pernicious hearsay evidence, or did the outrage stem from a philosophical antipathy over the Crown's refusal to produce the accusing witness to face Throckmorton? Policymakers' answers to this question will assist them in designing a hearsay rule that aligns with the policy preferences of their citizens and achieves lasting popular legitimacy.

184. Convergent validity is defined as the ability to demonstrate an empirical phenomenon across a variety of populations and experimental designs. *See, e.g.*, Donald T. Campbell & Donald W. Fiske, *Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix*, 56 PSYCHOL. BULL. 81, 100 (1959).