

Seeing is Believing: The Anti-Inference Bias

EYAL ZAMIR, ILANA RITOV, AND DORON TEICHMAN*

A large body of studies suggests that people are reluctant to impose liability on the basis of circumstantial evidence alone, even when this evidence is more reliable than direct evidence. Current explanations for this pattern of behavior focus on factors such as the tendency of fact finders to assign low subjective probabilities to circumstantial evidence, the statistical nature of such evidence, and the fact that direct evidence can rule out with greater ease any competing factual theory regarding liability.

This Article describes a set of four new experiments demonstrating that even when these factors are controlled for, the disinclination to impose liability based on indirect evidence remains. While these findings do not necessarily refute the existing theories, they indicate that these theories are incomplete and point to the existence of a deep-seated bias against basing liability on inferences—an anti-inference bias. The Article discusses the potential policy implications of the new findings for procedural and substantive legal norms.

INTRODUCTION	196
I. DOCTRINAL AND THEORETICAL BACKGROUND	197
II. EXPERIMENTAL FINDINGS	204
A. EXPERIMENT 1 (“HIGHWAY”): CONVICTION BASED ON TECHNOLOGICAL EVIDENCE	205
B. EXPERIMENT 2 (“ENFORCEMENT”): EX ANTE POLICY MAKING	207
C. EXPERIMENT 3 (“BUS”): CONVICTION BASED ON EYEWITNESS TESTIMONY	208
D. EXPERIMENT 4 (“ANTIBIOTICS”): FORENSIC EVIDENCE IN CIVIL LITIGATION	211

* Eyal Zamir is the Augusto Levi Professor of Commercial Law, Faculty of Law, Hebrew University of Jerusalem. Ilana Ritov is a Professor at the School of Education and Center for Rationality, Hebrew University of Jerusalem. Doron Teichman is the Joseph H. and Belle R. Braun Professor, Faculty of Law, Hebrew University of Jerusalem and the Justin D’Atri Visiting Professor of Law, Business & Society, Columbia Law School (Fall 2013). We would like to thank Netta Barak-Corren, Christoph Engel, Oren Gazal-Ayal, Werner Güth, Ofer Malcai, Amit Pundik, Amos Schurr, Alex Stein, Joshua Teitelbaum, Avishalom Tor, Fredrick Vars, Kathryn Zeiler, and the participants of workshops at the Georgetown University Law Center, William and Mary Law School, the Hebrew University of Jerusalem, the Max Planck Institute of Economics at Jena, the Center for Interdisciplinary Research at Bielefeld University, the University of Pennsylvania Law School, the Annual Conferences of the American and Israeli Law and Economics Associations, the Joint Conference of the Notre Dame Research Program on Law and Market Behavior and the University of Haifa Faculty of Law Aptowitz International Center for Risk, Liability, and Insurance and the Cornell-Tel Aviv Conference on Empirical Legal Studies, for valuable comments on earlier drafts. Excellent research assistance for this project was provided by Michael Cohen, Meirav Furth, and Ori Katz. This research was supported by the Israel Science Foundation (grants No. 100/11 and 1283/11) and the Aharon Barak Center for Interdisciplinary Legal Research.

III. DISCUSSION AND POLICY IMPLICATIONS.....	215
A. GENERAL DISCUSSION.....	215
B. POLICY IMPLICATIONS.....	221
C. LIMITATIONS AND FUTURE RESEARCH.....	227
IV. CONCLUSION.....	229

INTRODUCTION

Sherlock Holmes, Hercules Poirot, Sir Henry Merrivale, and even Dr. Gregory House have long been admired for their ability to solve mysteries using logic, deductive reasoning, and inferences from circumstantial evidence. However, the creators of detective stories and TV dramas almost never rely solely on this type of evidence. The detective story or TV episode will typically include a climactic scene in which direct evidence, such as a confession by the culprit, is presented. This serves to reassure the readers and viewers that the inferential arguments were indeed correct. The implicit assumption seems to be that inferences are insufficient to determine the culpability of the suspect (an assumption which is reinforced by the technique of incorporating misleading circumstantial evidence into detective stories and TV dramas).

The same is plausibly true in the courtroom. There are grounds to believe that judicial fact finders treat direct and circumstantial evidence differently and are more reluctant to impose liability on the basis of circumstantial evidence alone.¹ In exhibiting this reluctance, fact finders sometimes follow legal norms that caution against relying on circumstantial evidence, but more often disregard legal norms that deny the relevance of this distinction.² Scholars largely agree that this tendency is indefensible.³ In fact, it is sometimes noted that circumstantial evidence is *more* reliable than direct evidence, and thus discounting its probative value is a “paradox.”⁴

Several explanations have been offered for the tendency not to impose liability solely on the basis of circumstantial evidence. Most of these explanations, however, are incomplete at best.⁵ They are incomplete because too often they

1. See Kevin Jon Heller, *The Cognitive Psychology of Circumstantial Evidence*, 105 MICH. L. REV. 241, 247–55 (2006).

2. See Richard K. Greenstein, *Determining Facts: The Myth of Direct Evidence*, 45 HOUS. L. REV. 1801, 1802–04 (2009) (surveying American federal and state case law); *infra* Part II.

3. See 1A JOHN HENRY WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW 961 (Peter Tillers rev. ed., 1983) (“Wigmore’s view that circumstantial evidence may be as persuasive and as compelling as testimonial evidence, and sometimes more so, is now generally accepted.”); Greenstein, *supra* note 2, at 1804 (stating that “common beliefs about the significance of the distinction [between direct and circumstantial evidence] are false”).

4. Heller, *supra* note 1, at 244; see also Greenstein, *supra* note 2, at 1803 (pointing to the incompatibility between instructions that caution against relying on circumstantial evidence and the empirical data “indicating that at least some types of circumstantial evidence are actually *more* reliable than familiar categories of direct evidence” (emphasis in original)).

5. For a complete review, see *infra* Part II.

conflate the direct versus circumstantial evidence distinction with related ones, such as between eyewitness testimonies and scientific data, between concrete and statistical proof, and between probabilistic and conclusive evidence. Once we recognize that circumstantial evidence may be either eyewitness or forensic, concrete or statistical, probabilistic or conclusive, and so forth, we can appreciate the limitations of theories that rest on these distinctions.

This Article describes a series of new experiments demonstrating that even when we have controlled for such factors as the eyewitness/forensic, concrete/statistical, and probabilistic/conclusive character of the evidence, the disinclination to impose liability based on non-direct evidence remains. While our findings do not necessarily refute the existing theories, they indicate that these theories are insufficient and point to the existence of a deep-seated bias against basing liability on inferences from circumstantial evidence—an anti-inference bias.

The existence of an anti-inference bias gives rise to numerous complex normative questions that have not been fully addressed by the legal literature. On one hand, like other biases and heuristics, the anti-inference heuristic very often yields accurate decisions and might be a reasonable rule of thumb for people to follow in making their daily choices (that normally do not involve a courtroom). This suggests that debiasing judicial fact finders might require them to reach outcomes that are not aligned with the intuitions of most people regarding the desirable outcome of trial—a possibility that raises both practical and normative concerns. On the other hand, the anti-inference bias might give rise to systematic errors, especially in the context of judicial fact finding. This possibility highlights the need to design new and sophisticated tools that will help debias fact finders in order to improve the quality of their choices. Alternatively, to the extent the bias generates a hurdle to assigning legal responsibility in situations that require fact finders to draw inferences, policy makers can overcome it by introducing legal presumptions and by broadening the scope of legal liability. For example, in the context of the law of criminal attempts, legislators could deal with the reluctance of fact finders to deduce liability in cases of incomplete attempts (i.e., cases involving the preparation-attempt boundary) by criminalizing preparatory acts. Once such acts are criminalized, fact finders are no longer asked to infer whether an attempt at a greater crime took place, and the anti-inference bias no longer plays a role.

The Article is structured as follows. Following this brief introduction, Part I defines circumstantial evidence and surveys the extant explanations for the persistent reluctance to base criminal and civil liability on such evidence. Part II presents four new experiments designed to inquire whether this reluctance is unique to statistical, probabilistic, inconclusive, forensic, or non-case-specific evidence. Finally, Part III analyzes the contribution of our results to better understanding the behavior of fact finders, and discusses the potential normative implications of our findings.

I. DOCTRINAL AND THEORETICAL BACKGROUND

Direct evidence, if accepted, proves a material fact without the mediation of a deductive process. Circumstantial evidence, in contrast, even if accepted, requires

an additional mental step of inference in order to determine whether the material fact did or did not exist.⁶ The very existence of this distinction is, however, contested analytically. Some commentators argue that any determination of facts by judges or juries—including determination based on eyewitness testimony—inevitably involves inferences.⁷ For example, from the fact that the witness testifies that she saw the accused stabbing the victim, and the fact finder's impression that the witness is telling the truth, the fact finder infers that the accused actually stabbed the victim. According to this account, direct evidence is a myth.⁸ Direct evidence may thus be described as requiring a single inference (from the evidence to the material fact), while circumstantial evidence requires at least two: from the evidence to an underlying fact and from the underlying fact to the material one.

Under any of these descriptions, elements of what is conventionally viewed as direct evidence are sometimes circumstantial. For instance, when a litigant takes the stand and testifies that she made an oral agreement with the defendant on a certain date, her testimony would ordinarily be viewed as direct evidence. However, typically much of this testimony will refer to the prior relationships between the parties and their subsequent behavior in order to corroborate the litigant's assertion that the alleged agreement was actually made. These parts of her testimony are analytically circumstantial evidence.⁹ One may thus distinguish between evidence that is perceived as wholly circumstantial and evidence that is not. In addition, to some extent the distinction between the two types of evidence is a matter of degree: Some evidence requires more inferential steps than other evidence to draw a conclusion about the material fact. Often, judicial fact finding consists of several (possibly multistage) inferences, some independent and some interconnected, based on a multiplicity of evidence.¹⁰

These complexities notwithstanding, the distinction between direct and circumstantial evidence is entrenched in popular and legal vocabulary. However analytically described or critiqued, this distinction has bearing on one's perceptions and decisions. Inasmuch as people treat circumstantial evidence differently than direct evidence, this differentiation may have considerable ramifications in both

6. 1 CHARLES T. MCCORMICK, MCCORMICK ON EVIDENCE 308 (Kenneth S. Broun et al. eds., 6th, one-vol. ed. 2006); ALBERT J. MOORE, PAUL BERGMAN & DAVID A. BINDER, TRIAL ADVOCACY: INFERENCES, ARGUMENTS, AND TECHNIQUES 2–3 (1996); PETER MURPHY, MURPHY ON EVIDENCE 20–21 (10th ed. 2008).

7. See, e.g., WIGMORE, *supra* note 3, at 952–56.

8. See Greenstein, *supra* note 2. This position echoes the reductionist view in the philosophical debate over the epistemological problems of testimony. According to this view, a testimony is not an autonomous source of epistemic authority; rather, its reliability rests on additional sources of knowledge. See generally C.A.J. COADY, TESTIMONY: A PHILOSOPHICAL STUDY (1992); Jonathan Adler, *Epistemological Problems of Testimony*, STAN. ENCYCLOPEDIA PHIL. (Aug. 27, 2012), <http://plato.stanford.edu/entries/testimony-episprob>.

9. MOORE ET AL., *supra* note 6, at 3–4; Paul Bergman, *A Bunch of Circumstantial Evidence*, 30 U.S.F. L. REV. 985, 988–89 (1996).

10. DAVID A. SCHUM, EVIDENCE AND INFERENCE FOR THE INTELLIGENCE ANALYST *passim* (1987) (using symbols and diagrams to describe complex inferential processes).

criminal and civil litigation, as often there is no direct evidence bearing on the disputed facts.

As a doctrinal matter, some legal systems instruct, or used to instruct, fact finders to avoid convictions based exclusively on circumstantial evidence.¹¹ Other legal systems have allowed fact finders to rely on such evidence, yet have required them to exercise special caution in both criminal and civil cases.¹² At times, stricter limitations have been imposed on drawing an inference upon an inference—namely, inferring the existence or nonexistence of a material fact from another fact that was itself inferred from other evidence. According to one version of this limitation, even in civil litigation, an inference upon an inference is impermissible, and according to another version it is only permitted if the first inference meets the “beyond a reasonable doubt” standard.¹³

Clearly, the modern trend is to abandon rules limiting the use of circumstantial evidence, including an inference upon an inference.¹⁴ Instead, fact finders are instructed to determine the reliability and weight of any evidence, direct or indirect, without prejudice against the latter.¹⁵ Vestiges of the distrust of circumstantial evidence can nevertheless be found in jury instructions in some jurisdictions. Thus, for instance, the Mississippi Model Jury Instructions provide as follows:

The court instructs the jury that if the State has relied on circumstantial evidence to establish its theory of guilt of the defendant, then the evidence for the State must be so strong as to establish the guilt of the defendant, not only beyond a reasonable doubt, but the evidence must

11. See, e.g., Irene Merker Rosenberg & Yale L. Rosenberg, “*Perhaps What Ye Say Is Based Only on Conjecture*”—Circumstantial Evidence, Then and Now, 31 HOUS. L. REV. 1371, 1376–90 (1995) (discussing Jewish law).

12. See, e.g., *id.* at 1390–402 (analyzing 19th century and current American law).

13. L. JONATHAN COHEN, THE PROBABLE AND THE PROVABLE 68–73 (1977) (discussing the difficulty in relying on inference upon inference); RICHARD EGGLESTON, EVIDENCE, PROOF AND PROBABILITY 35, 39–40, 237–40 (2d ed. 1983); WIGMORE, *supra* note 3, at 1106–19 (critically surveying the case law); Michael Foster, *A Review and Reconsideration of Florida’s Rule Against Basing an Inference on an Inference in Civil Cases*, 23 STETSON L. REV. 743 (1994) (critically discussing the rule as applied in Florida).

14. See, e.g., EGGLESTON, *supra* note 13, at 35, 39–40, 237–40 (rejecting Cohen’s account of the law). On the roots of this trend in early modern English law, see BARBARA J. SHAPIRO, “BEYOND REASONABLE DOUBT” AND “PROBABLE CAUSE”: HISTORICAL PERSPECTIVES ON THE ANGLO-AMERICAN LAW OF EVIDENCE 200–43 (1991).

15. See, e.g., *Holland v. United States*, 348 U.S. 121, 140 (1954) (stating that, in terms of reliability, “[c]ircumstantial evidence . . . is intrinsically no different from testimonial evidence”); SIDNEY L. PHIPSON, PHIPSON ON EVIDENCE 5 (Hodge M. Malek et al. eds., Thomson Reuters 17th ed. 2010) (1892) (“Little is to be gained from a comparison of [direct and indirect evidence’s] weight, since . . . both forms admit of every degree of cogency from the lowest to the highest.”); WIGMORE, *supra* note 3, at 957–64, 1120–38 (surveying cases in which judges have repudiated the rule against basing an inference on an inference); Rosenberg & Rosenberg, *supra* note 11, at 1400–02 (showing that most courts in the United States follow the *Holland* dictum).

be so strong as to exclude every other reasonable hypothesis other than that of guilt.¹⁶

Similar instructions may be found in case law.¹⁷ In the same vein, while the rule restricting an inference upon an inference has largely been abandoned, it is still in force in some jurisdictions.¹⁸

Commentators generally agree that the prejudice against circumstantial evidence is unwarranted.¹⁹ Yet, empirical and experimental studies have long demonstrated that fact finders are far more inclined to base their conclusions on direct evidence.²⁰ Fact finders sometimes tend to undervalue the reliability and probative value of circumstantial evidence, including statistical and forensic evidence,²¹ and to overvalue the trustworthiness and weight of direct evidence, such as eyewitness testimonies.²²

16. MISSISSIPPI MODEL JURY INSTRUCTIONS—CRIMINAL § 1:18 (Miss. Judicial Coll. 2012). See also Carl N. Hammarskjold, *Smokes, Candy, and the Bloody Sword: How Classifying Jailhouse Snitch Testimony as Direct, Rather than Circumstantial, Evidence Contributes to Wrongful Convictions*, 45 U.S.F. L. REV. 1103, 1120–22 (2011) (describing the Mississippi instructions in a historical context).

17. See, e.g., Rosenberg & Rosenberg, *supra* note 11, at 1399 (describing state courts that refused to follow the *Holland* ruling); see also Phipson, *supra* note 15, at 414 (citing approvingly the Judicial Studies Board Specimen Direction that calls for examining circumstantial evidence with care).

18. See, e.g., *Stanley v. Marceaux*, 991 So. 2d 938, 940 (Fla. Dist. Ct. App. 2008) (quoting *Nielsen v. City of Sarasota*, 117 So. 2d 731, 733 (Fla. 1960) (discussing the rules governing the use of circumstantial evidence in civil cases); Foster, *supra* note 13 (critically analyzing the law in Florida).

19. See, e.g., Phipson, *supra* note 15, at 5; Wigmore, *supra* note 3, at 957–64; Greenstein, *supra* note 2, at 1804 (questioning the very distinction between direct and circumstantial evidence). But see Rosenberg & Rosenberg, *supra* note 11, at 1402–22 (criticizing the equation of the probative weight of circumstantial and direct evidence).

20. See Heller, *supra* note 1, at 250–52 (contending that “jurors consistently underestimate the probative value of [circumstantial forensic evidence]”).

21. Jane Goodman, *Jurors’ Comprehension and Assessment of Probabilistic Evidence*, 16 AM. J. TRIAL ADVOC. 361, 375 (1992) (describing experimental findings indicating that “where the probabilistic evidence is most incriminating and most probative, jurors may tend to underuse or undervalue the evidence”); Dale A. Nance & Scott B. Morris, *Juror Understanding of DNA Evidence: An Empirical Assessment of Presentation Formats for Trace Evidence with a Relatively Small Random-Match Probability*, 34 J. LEGAL STUD. 395, 407–18 (2005) (reporting the results of a large-scale empirical study indicating that jurors tend to undervalue forensic match evidence).

22. See, e.g., Heller, *supra* note 1, at 247–50, 253–54 (surveying the psychological literature and the data on false convictions); R.C.L. Lindsay, Gary L. Wells & Carolyn M. Rumpel, *Can People Detect Eyewitness-Identification Accuracy Within and Across Situations?*, 66 J. APPLIED PSYCHOL. 79, 80–86 (1981) (reporting experimental results demonstrating that subjects who viewed the cross-examination of eyewitnesses were unable to distinguish between accurate and false identifications); Steven Penrod & Brian Cutler, *Witness Confidence and Witness Accuracy: Assessing Their Forensic Relation*, 1 PSYCHOL. PUB. POL’Y & L. 817, 817 (1995) (“Jurors overbelieve eyewitnesses, have difficulty reliably differentiating accurate from inaccurate eyewitnesses, and are not adequately sensitive to aspects of witnessing and identification conditions.”).

Several explanations have been offered for this phenomenon. One is that, unlike direct evidence, inferences are based on generalizations, “which by definition are accurate less than 100% of the time.”²³ Even reliable circumstantial evidence may be compatible with numerous competing inferences; hence, it is necessarily “polyvocal,” to use the term coined by Kevin Heller.²⁴ Unlike the first explanation, the second does not assume that circumstantial evidence is objectively less conclusive than direct evidence, yet it argues that circumstantial evidence may be subjectively perceived as such. Focusing on people’s reluctance to use statistical evidence as a basis for imposing liability, it has been argued that fact finders make decisions on the basis of subjective probability estimates, which may differ from the objective probability estimates.²⁵ Fact finders may believe that the reliability and probative weight of circumstantial evidence is considerably lower than it actually is, and therefore be reluctant to rely on it (the opposite would hold for direct evidence).

A third explanation focuses on the typical features of eyewitness testimonies, often atypical of circumstantial evidence. Among other things, eyewitness testimonies provide “a verbal representation of the crime itself,” whereas indirect evidence is often abstract.²⁶ Direct evidence is a story-like narrative, while arguments based on circumstantial evidence often resemble deductive reasoning.²⁷ The former are often vivid, concrete, and stirring, while the latter are pallid, general, and unexciting.²⁸ These characteristics make it easier for fact finders to form a coherent story of the events from direct evidence and, as a result, more likely to convict.²⁹

A fourth explanation, resonating with both the alleged generalization-based nature of circumstantial evidence (underlying the first explanation) and its nonconcrete quality (part of the third explanation), has to do with the case-specificity of the evidence. A central justification for the reluctance to accept a claim on the basis of naked statistical evidence draws on the distinction between the probability that a certain fact is true and the weight or resiliency of the evidence supporting the claim. Decision makers may sensibly reject a claim even if the probability that the plaintiff’s version is correct meets the controlling standard of proof, if assessment of this probability rests on too little information or on general,

23. Bergman, *supra* note 9, at 988.

24. Heller, *supra* note 1, at 267–68; *see also* MOORE ET AL., *supra* note 6, at 4–7 (arguing that “circumstantial evidence frequently gives rise to multiple and even conflicting inferences”).

25. *See* William C. Thompson & Edward L. Schumann, *Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor’s Fallacy and the Defense Attorney’s Fallacy*, 11 LAW & HUM. BEHAV. 167, 172–81 (1987) (experimentally demonstrating subjects’ errors in using probabilistic information); Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1344–50 (1971).

26. Heller, *supra* note 1, at 265.

27. *Id.* at 265–66.

28. *Id.* at 267–80.

29. *Cf.* Nicholas Scurich & Richard S. John, *Trawling Genetic Databases: When a DNA Match Is Just a Naked Statistic*, 8 J. EMPIRICAL LEGAL STUD. 49, 58–59 (2011) (discussing the “story model of juror decision making”).

non-case-specific evidence.³⁰ For example, the fact that there were 1000 spectators in a rodeo but only 400 tickets had been sold means that for any spectator chosen at random, there is a 60% probability of her being a gatecrasher. The fact finder may nevertheless dismiss a claim based on this evidence alone as resting on too thin an evidentiary basis.³¹

Yet another possibility is that from the perspective of the fact finders, decisions based on direct testimonies involve a smaller degree of responsibility. If it turns out that a testimony was inaccurate or deceptive, the fact finder can rationalize that the witness is to blame for the erroneous verdict. In contrast, if it transpires that liability was erroneously imposed due to a faulty inference from circumstantial evidence, the responsibility arguably lies with the fact finder.³² To avoid a feeling of regret, a fact finder would likely be disinclined to rely on circumstantial evidence.³³ This explanation bears some similarity to the epistemological notion that when a speaker makes testimonial assertions, she undertakes responsibility for the accuracy of what she says; and when a hearer takes a speaker's word for something, she ascribes her authority.³⁴

The sixth and seventh theories were originally offered in the context of naked statistical evidence and later extended by Kevin Heller to circumstantial evidence more generally.³⁵ In a seminal study, Gary Wells conjectured that "in order for evidence to have a significant impact on people's verdict preferences, one's hypothetical belief about the ultimate fact must affect one's belief about the evidence."³⁶ For example, the mere fact that 80% of the buses in a certain town

30. For detailed discussions of this claim, primarily in the context of naked statistical evidence, see COHEN, *supra* note 13, at 36–39; ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW 40–56, 80–106 (2005); L. Jonathan Cohen, *The Role of Evidential Weight in Criminal Proof*, 66 B.U. L. REV. 635 (1986); D.H. Kaye, *Apples and Oranges: Confidence Coefficients and the Burden of Persuasion*, 73 CORNELL L. REV. 54 (1987); Amit Pundik, *What Is Wrong with Statistical Evidence? The Attempts to Establish an Epistemic Deficiency*, 27 CIV. JUST. Q. 461, 474–87 (2008) (criticizing the weight argument).

31. COHEN, *supra* note 13, at 74–76 (discussing the gatecrasher paradox); L. Jonathan Cohen, *Subjective Probability and the Paradox of the Gatecrasher*, 1981 ARIZ. ST. L.J. 627, 627 (replying to reactions to his original analysis).

32. Heller, *supra* note 1, at 287.

33. For the claim that anticipation of possible ex post regret affects people's decisions ex ante, see generally David E. Bell, *Regret in Decision Making Under Uncertainty*, 30 OPERATIONS RES. 961 (1982); Richard P. Larrick & Terry L. Boles, *Avoiding Regret in Decisions with Feedback: A Negotiation Example*, 63 ORG. BEHAV. & HUM. DECISION PROCESSES 87 (1995); Graham Loomes, *Further Evidence of the Impact of Regret and Disappointment in Choice Under Uncertainty*, 55 ECONOMICA 47 (1988); Graham Loomes & Robert Sugden, *Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty*, 92 ECON. J. 805, 809 (1982).

34. Adler, *supra* note 8 (surveying the philosophical literature); see also Arnon Keren, *Epistemic Authority, Testimony and the Transmission of Knowledge*, 4 EPISTEME 368, 368 (2007) (arguing that "to trust a speaker is to grant her epistemic authority on the asserted proposition, and hence to see her opinion as issuing a second order, preemptive reason for believing the proposition").

35. Heller, *supra* note 1.

36. Gary L. Wells, *Naked Statistical Evidence of Liability: Is Subjective Probability Enough?*, 62 J. PERSONALITY & SOC. PSYCHOL. 739, 746 (1992) (quote italicized in original).

belong to the blue bus company and 20% to the grey bus company is insufficient to find the former liable for an accident caused by an unidentified bus, because the determination of liability would not change one's belief about the accuracy of the statistical fact. The statistical data remains true whether or not a blue bus was involved in the accident. In contrast, when a weigh station attendant testifies that according to his records, a blue bus was weighed in the nearby station just before the accident—thus tying that bus company to the accident—the determination of liability would more likely bear on the reliability of this testimony, even if the defendant had already established that the records were wrong 20% of the time. In this case, determining which bus was involved in the accident does bear on the accuracy of the weigh station's records and the reliability of the attendant's testimony.

The seventh and last explanation is inspired by Daniel Kahneman and Amos Tversky's simulation heuristic. Kahneman and Tversky argued that the ease with which different scenarios can be simulated is used to judge the probability of specific events, to assess the causal connection between two events, and so forth.³⁷ In the present context, Keith Niedermeier and his coauthors argued that the willingness to ground liability on statistical evidence depends on how easily one can imagine an alternative scenario that would be compatible with the evidence.³⁸ Arguably, it is easier to imagine such a scenario when the only evidence is circumstantial, since by its very nature, such evidence does not prove the material

37. Daniel Kahneman & Amos Tversky, *The Simulation Heuristic*, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 201, 201–03 (Daniel Kahneman, Paul Slovic & Amos Tversky eds., 1982).

38. Keith E. Niedermeier, Norbert L. Kerr & Lawrence A. Messé, *Jurors' Use of Naked Statistical Evidence: Exploring Bases and Implications of the Wells Effect*, 76 J. PERSONALITY & SOC. PSYCHOL. 533, 536–37 (1999); Scurich & John, *supra* note 29, at 58 (discussing Niedermeier and his coauthors' theory). Following the experimental design of Wells, *supra* note 36, Niedermeier and his coauthors presented the participants in their study with various vignettes describing a legal suit for damages. The lawsuit was filed by a woman whose dog was killed by a bus. In all the vignettes, there was an objective probability of 80% that the defendant—one of the only two bus companies in the county—was responsible for the accident. In the “complete match” version, the evidence was based on a comparison between the tracks left by the bus that ran over the dog and the tracks of all the buses belonging to the two companies. Each company operated ten buses, and the track prints matched eight of the defendant's buses and two buses belonging to the other company. In the “partial match” version, there was a partial match with only one of the defendant's buses and one of the other company's buses, indicating that there was an 80% probability that the defendant's bus was the one that ran over the dog and a 20% chance that it was a bus belonging to the other company. The responders were asked to indicate whether they would accept the claim against the defendant—the blue bus company—and to assess the probability that the defendant's bus was involved in the accident. While the assessed probability was similar under the two conditions, the willingness to accept the claim was much lower in the complete-match condition, arguably because in that case it was easier to imagine that the dog had been killed by one of the other company's two buses whose tracks matched those left by the bus that had run over the dog. *Id.*; cf. Heller, *supra* note 1, at 258–303 (extending the ease-of-simulation explanation and applying it to circumstantial evidence generally).

fact itself.³⁹ Hence, fact finders are considerably less willing to find defendants liable on the basis of such evidence. This explanation is associated with the third explanation discussed above, as direct evidence helps the fact finder to form a coherent story of the events, which then makes it harder to imagine an alternative story. In the same vein, Deanna Sykes and Joel Johnson argued that since comprehension of a witness's assertion entails an initial belief in that assertion, undoing the mental representation created by the testimony and imagining an alternative scenario requires cognitive effort.⁴⁰ Such an effort is not required when decision makers are presented with probabilistic evidence that does not include a concrete assertion that needs to be undone.⁴¹

All of these explanations can be—and have been—challenged. Thus, for example, the generality of the subjective probability argument (the second explanation) was called into question by Wells, who instead offered the fact-to-evidence theory.⁴² Wells's theory was in turn challenged by Niedermeier, Kerr, and Messé and Sykes and Johnson, who proposed the ease-of-simulation explanation.⁴³ The ease-of-simulation explanation was subsequently called into question by Arkes et al.⁴⁴ As the experiments in the next Part demonstrate, none of these arguments captures the full scope of the bias against indirect evidence.

II. EXPERIMENTAL FINDINGS

To test the scope of fact finders' aversion towards circumstantial evidence, we conducted a series of four experimental surveys using situational vignettes on a student population. In three of the four studies we employed a between-subject design. That is, participants were randomly assigned to one of several equally sized groups and given a questionnaire structured around a hypothetical case. Whereas half of them read a case in which the evidence described was direct in nature, the other half read a similar case in which the evidence was indirect. Participants were then asked a series of questions focusing on their willingness to assign legal liability based on the described evidence. In order to broaden the methodological robustness of our studies, in one of them we employed a within- (rather than between-) subject design. According to this framework all participants were presented with both types of evidence and were asked to choose between them.

39. Cf. Heller, *supra* note 1, at 258–66 (describing the different applications of the ease-of-simulation model to direct and circumstantial evidence).

40. Deanna L. Sykes & Joel T. Johnson, *Probabilistic Evidence Versus the Representation of an Event: The Curious Case of Mrs. Prob's Dog*, 21 BASIC & APPLIED SOC. PSYCHOL. 199 (1999).

41. *Id.* at 209–10.

42. Wells, *supra* note 36, at 742, 746.

43. Niedermeier et al., *supra* note 38, at 534, 536–37; Sykes & Johnson, *supra* note 40, at 200–01.

44. Hal R. Arkes, Brittany Shoots-Reinhard & Ryan S. Mayes, *Disjunction Between Probability and Verdict in Juror Decision Making*, 25 J. BEHAV. DECISION MAKING 276, 277–78 (2012).

A. Experiment 1 (“Highway”): Conviction Based on Technological Evidence

Since much of the existing literature on circumstantial evidence focuses on criminal liability, Experiment 1 was designed to examine how the type of evidence impacts the willingness to impose such liability—namely, convicting a driver of a traffic violation. Specifically, the experiment examined the difference between evidence that directly establishes the defendant’s fault and evidence that establishes fault indirectly, requiring an additional inferential step. The experiment held constant the nature of the evidence involved—both scenarios alluded to technological evidence. Furthermore, it controlled for the objective reliability of the evidence by supplying subjects with information regarding the error rates of the enforcement technology.

We used a between-subject design, assigning each participant randomly to one of two versions of a short scenario. The direct-evidence condition read as follows:

Speed cameras were installed on a toll road. The probability of an error in the camera system is 2%. The speed limit on this road is 100 KPH. According to the camera, a driver drove his car at a speed of 125 KPH late at night.

The inference condition was as follows:

Cameras that document the exact time at which each vehicle passes by them were installed on a toll road. The cameras do not document the speed of the passing vehicle, but from the distance between the cameras and the time that elapses between the points they document, it is possible to infer the driver’s speed in that section of the road. The probability of an error in the camera system is 2%. The speed limit on this road is 100 KPH. According to the time elapsed and the distance between the two cameras, a driver drove his car at a speed of 125 KPH late at night.

Note that while the enforcement technology in the two conditions differed, the text describing the objective probability of error was identical irrespective of how many cameras were employed (i.e., “[t]he probability of an error in the camera system is 2%”).

After reading the scenario, responders were reminded that the standard of proof in such cases is “beyond a reasonable doubt.” Based on this evidentiary threshold they were then asked to indicate whether they would convict the driver of violating the speed limit. The responses to this dichotomous question constituted the main dependent variable of the experiment.⁴⁵ In addition, in order to probe deeper into participants’ decision-making processes, we asked them two additional follow-up questions. Respondents who decided affirmatively with respect to the liability question were asked to indicate the penalty they would impose on the driver. The penalty could range from 50 NIS to 1000 NIS (1 NIS roughly equals \$0.25). Finally, all responders were asked to rate the fairness of conviction on the basis of the evidence presented, using a scale of 0 to 100, where 0 indicates “absolutely unfair” and 100 “absolutely fair.”

45. “Dependent variable” refers to the variables that are explained by the model.

The two questionnaires were posted on a website designed for controlled judgment and decision-making experiments. Responders were recruited by an e-mail message sent to students at the Faculty of Law of the Hebrew University of Jerusalem. To encourage participation, one out of ten participants was randomly selected to win a prize of 100 NIS (approximately \$25). A total of 88 law students (57 women, 31 men) who had finished their first year of studies participated in the study.⁴⁶ These students ranged in age from 19 to 40, with a mean of 24.

The dependent variable—rate of convictions—differed significantly by condition.⁴⁷ Whereas 81.4% of the responders in the direct-evidence condition decided to convict the driver, only 60% of the responders in the inference condition did.⁴⁸ Even though both groups faced the same objective probability, and even though both groups faced a similar situation involving a technological enforcement mechanism, respondents were far more likely to convict when supplied with direct evidence.⁴⁹

A more nuanced picture emerged when two follow-up questions were considered. Counting acquittals as a fine of zero, the average fine in the direct evidence group was 300.3 NIS, and the average fine in the inference group was 286.2 NIS. These averages, however, were not significantly affected by condition.⁵⁰ The mean fairness rating of convicting the driver in the direct-evidence condition (74.8 on a scale of 0 to 100, where 0 is “absolutely unfair” and 100 “absolutely fair”) was higher than in the inference condition (66.1 on the same scale). Nonetheless, this difference was only marginally statistically significant.⁵¹ The willingness to convict in the two conditions was correlated with the judged fairness of conviction, so that controlling for the judgment of fairness eliminated the effect of condition (direct vs. inference evidence)

46. In Israel, law is studied at the undergraduate level, with students earning an LLB.

47. A significant effect in a statistical measurement refers to the odds that a certain result was created by chance. In the context of this paper, every time a difference or a result is presented as “significant,” it means that there is less than a 5% likelihood that this difference was coincidental. The 5% level of significance is a common threshold used in statistical analysis in the social sciences. A 6% to 10% level of significance is commonly considered marginally significant. In technical terms results are significant when the p-value associated with them is 0.05 or below.

48. Chi-Square=4.835, $p=0.028$. A chi-square procedure can be used to test the hypothesis that two different outcomes were obtained by chance. Here we use it to test the conviction rates in the two conditions. Specifically, we aim to reject the hypothesis that the difference in the two rates arose due to chance.

49. Both gender and age had no statistically significant main effects and no significant interactions with the experimental conditions (in this or in any of the following experiments).

50. $p=0.832$. Similarly, the average fine excluding the acquittals did not significantly differ between the two conditions (368.9 vs. 477.0 for direct evidence and inference, respectively, $t(60)=1.472$, $p=0.15$).

51. $t(86)=1.64$, $p=0.1$.

on conviction rate.⁵² However, as will be discussed below, it is unclear what one can make of this finding.⁵³

B. Experiment 2 (“Enforcement”): Ex Ante Policy Making

Experiment 2 set out to deepen our understanding of the disinclination to base legal liability on indirect evidence by exploring whether the results obtained in the first experiment, which focused on an ex post judicial decision, would be replicated with respect to an ex ante choice of an enforcement system. The shift from an ex post to an ex ante perspective can prove helpful from both substantive and methodological perspectives. From a substantive perspective, as noted above, one explanation of the disinclination to base liability on circumstantial evidence alludes to the issue of the responsibility of the fact finder for an erroneous judicial outcome when she wrongfully deduces liability.⁵⁴ Once the decision is shifted to the policy context, it is difficult to see any true difference between direct and circumstantial evidence regarding the responsibility of the policy maker. From a methodological perspective, the move to a policy question enabled us to employ an alternative experimental design. Rather than assigning participants randomly to one of two conditions and comparing the two, we employed a within-subject design in which subjects revealed their preferences regarding the different types of evidence. Earlier research suggests that, for the same set of options, tasks that rest on isolated judgments and tasks that call for comparative judgment may yield different rankings.⁵⁵ Replicating the results of the first experiment (between-subject design) in a within-subject design (the second experiment) could thus bolster our findings.

Participants in Experiment 2 were instructed to imagine that they were members of a team of experts whose job is to set policy regarding the enforcement of traffic laws. They were then asked to choose between two speed-enforcement systems. One system was based on speed-enforcement cameras, as in the direct-evidence condition of Experiment 1. The other system was based on two cameras documenting the exact time at which a car passes by, from which it is possible to infer the driver’s speed, as in the inference condition of Experiment 1. The cost of purchasing, installing, and operating the two systems was identical, and the accuracy of both systems was 99%. In order to avoid any potential effect that might be caused by the label attached to each system, we simply named them A and B. Furthermore, to assure that decisions were not influenced by the order in which

52. Predicting conviction decision by a logistic regression analysis that includes both the condition (direct evidence vs. inference as a dummy variable, coded as 1 for direct evidence and 0 for inference condition) and fairness judgments as independent variables renders the effect of the condition nonsignificant ($B=.782$, $Wald=1.422$, $p=.233$ and $B=.072$, $Wald=20.696$, $p<.001$ for condition and fairness, respectively).

53. See *infra* notes 97–99 and accompanying text.

54. See *supra* notes 32–34 and accompanying text.

55. See generally Christopher K. Hsee, George F. Loewenstein, Sally Blount & Max H. Bazerman, *Preference Reversals Between Joint and Separate Evaluations of Options: A Review and Theoretical Analysis*, 125 PSYCHOL. BULL. 576 (1999) (reviewing the literature and proposing an explanation for the phenomenon of preference reversal between joint and separate evaluations).

participants read about the two systems, the order was varied randomly between subjects.

Responders were given three options: preference for system A, preference for system B, and indifference. Responders with a preference for one of the two systems were then asked whether they would stick with their choice even if it was 98% accurate and the other system remained 99% accurate. Finally, they were asked to indicate the lowest level of accuracy at which they would still choose their preferred system over the alternative, assuming the other remained 99% accurate.

Participants in Experiment 2 were recruited voluntarily in one of their law classes. A total of 49 first-year students participated in the study (31 women, 18 men). Their ages ranged from 18 to 32, with a mean of 23.

The results of the experiment confirmed the existence of a bias against circumstantial evidence. Of the 49 responders, 53.1% preferred the “direct-evidence system,” 24.5% preferred the “inference system,” and 22.4% were indifferent. The percent of responders who preferred the direct-evidence system was significantly higher than the percent preferring the inference one.⁵⁶

Of the 26 responders who preferred the direct-evidence system, 11 would have preferred it even if it was 98% accurate and 13 would not (two did not answer this question). The mean lowest accuracy they would accept was 95.2%. Of the 12 responders who preferred the inference system, 8 would have preferred it even if it was 98% accurate and 4 would not. The mean lowest accuracy they would accept was 92.2%. Although these findings may suggest that subjects, who overcame the anti-inference bias and preferred the indirect-evidence enforcement regime, exhibited a stronger preference for this technology relative to the subjects who preferred the direct-evidence regime, neither the choice proportion nor the threshold for preference change were significantly different when comparing the two groups.⁵⁷

C. Experiment 3 (“Bus”): Conviction Based on Eyewitness Testimony

Experiment 3 was designed to further examine how general the findings of Experiments 1 and 2 were along several dimensions. First, we wanted to investigate whether the results were due to any unique feature of the pertinent enforcement mechanisms. For example, one could conjecture that responders were reluctant to convict in the inference condition because the two-camera system might have been perceived as more intrusive and a greater threat to people’s privacy. Thus, Experiment 3 was designed in a way that such potential issues would not arise.

Second, we wanted to see whether the results of Experiment 1 were a function of the technological nature of the evidence or whether they would hold for eyewitness testimony as well. Much of the literature on direct versus circumstantial evidence deals with the distinction between eyewitness—the paradigmatic direct

56. $z=2.953$, $p<0.01$. The order of presentation did not significantly affect preference (Chi-Square=1.414, $p=0.493$). Thus, we could treat all subjects as one group. Z-test is the standard statistical test for comparison of means or proportions.

57. More specifically, the p-values were $p=.231$ and $p=.386$ for the comparison of the choice proportions and the preference change threshold, respectively.

evidence—and forensic, indirect evidence.⁵⁸ To control for the explanatory force of this distinction, Experiment 1 held constant the technological nature of the evidence across both experimental conditions. Similarly, in Experiment 3 both the direct evidence group and the inference group were presented with eyewitness testimony.

Third, we wanted to test whether the results are driven by the responders' subjective probability estimates. While the design of Experiments 1 and 2 held the objective probability of a blameworthy behavior constant, subjects might have assigned different subjective probabilities notwithstanding.⁵⁹ In order to control for this somewhat elusive factor, we asked subjects not only to decide whether to convict the accused but also to assess the probability of guilt.

Finally, we were interested in examining whether the difference between the two conditions in Experiment 1 had to do with the perceived responsibility-taking of the decision maker.⁶⁰ Thus, instead of asking participants how they would decide the case, we asked them how a judge should decide the case. Presumably, shifting the act of making a decision to a third party should diminish the responsibility subjects feel towards their choices.

As in Experiment 1, this experiment used a between-subject design, employing two versions of a short scenario, followed by a few questions. The two versions described a tourist bus that was stuck in a sparsely populated area late in the evening. A policeman who arrived on the scene assisted the driver in arranging for minibuses to bring the tourists to their destination. In the direct-evidence condition, the policeman, who got on the bus, saw that there were fifty-four tourists on it, despite the fact that its permit allowed it to carry only fifty passengers. In the inference condition, two minibuses arrived empty, were completely filled up, and drove away. It then turned out that four tourists were left, and an additional vehicle was summoned for them. Since each minibus had seats for twenty-five passengers, the policeman inferred that the bus driver had violated the terms of the permit, which allowed only fifty passengers on the bus. In both scenarios, based on the policeman's report, the driver was accused of carrying too many passengers. The driver pleaded not guilty. The judge had the impression that the policeman was a reliable person and estimated the likelihood that he had miscounted the passengers as very low: one in twenty. As elsewhere, the exact same wording was used in both conditions to describe the possibility of error.⁶¹

58. For a review of the literature on this point, see *supra* notes 26–29, 32–34 and accompanying text.

59. For a review of the literature on this point, see *supra* note 25 and accompanying text.

60. For a review of the literature on this point, see *supra* notes 32–34 and accompanying text.

61. The vignette did not detail how the policeman knew that there were twenty-five passenger seats in each minibus, that the minibuses arrived empty, and that the minibuses were full when they set out. However, since these facts were presented as given, they should not have affected the reliability of the policeman's inference. At any rate, the reliability of the policeman's count, as assessed by the judge, was described identically in both conditions in order to eliminate any effect of such hypothetical doubts about the judged probability of the driver's guilt.

After being reminded of the standard of proof in criminal proceedings (“beyond a reasonable doubt”), participants were asked whether in their opinion the judge should find the driver guilty. Much like in the Experiment 1, the responses to this dichotomous question constituted the main dependent variable of the experiment. Following this question, participants were asked to assess the probability that the driver had actually transported more passengers than permitted. This second question allowed us to collect data regarding subjects’ subjective valuation of the probability of guilt. Finally, participants were asked to rate the fairness of convicting the driver on a scale of 1 to 9.

Participants in the third experiment were recruited voluntarily in one of their law classes. A total of 117 first-year law students (59 women, 58 men) answered the questionnaire during the first week of their studies. They ranged in age from 18 to 36, with a mean of 23. Each participant was randomly assigned to one of the two conditions.

The results of the third experiment confirmed the existence of an anti-inference bias yet again. Whereas 64.8% of the participants in the direct-evidence condition determined that the judge should convict the driver, only 38.1% of the participants in the inference condition were of this opinion. This result reflects a highly significant difference between the two groups.⁶²

Analyzing responders’ subjective valuations regarding liability sheds additional light on these findings. The participants estimated the probability of the driver actually violating the terms of the permit as considerably lower than the probability implied by the vignette. While the chance of miscounting according to the judge’s impression was described as one in twenty, implying a 95% reliability of the testimony, the mean estimated probability of the driver actually committing the violation was 80.1% across the two conditions. More importantly, the mean probability estimates were not significantly different for the two conditions: 82.2% and 78.3% for direct-evidence and inference conditions, respectively.⁶³

Furthermore, although—as one would expect—the probability estimate affected the conviction decision, it did not mediate the effect of the condition. A logistic regression analysis⁶⁴ predicting the conviction decision based on the condition subjects were assigned to (direct evidence vs. inference) and their subjectively estimated probability yielded significant effects of both probability and condition.⁶⁵ In other words, the tendency to convict remained significantly lower in the circumstantial evidence condition, even when controlling for the estimated probability of guilt. As in Experiment 1, the effect of the condition, direct evidence versus inference, on the conviction decisions was possibly mediated by the effect of this factor on judged fairness.⁶⁶ It is also possible, however, that the similar effects

62. Chi-Square=8.304, $p=0.004$.

63. $p=0.29$.

64. Regression analysis is a method that enables documenting the way in which each of the independent variables influences the value of the dependent variable.

65. $B=1.091$, $Wald=7.092$, $p<.01$ and $B=0.038$, $Wald=11.097$, $p<0.01$ for the effects of the condition and the subjective probability, respectively.

66. Convicting the driver was perceived as fairer in the direct-evidence condition than in the inference condition: mean ratings of 6.09 and 4.70 for direct evidence and inference, respectively, on a scale of 1 to 9, where 1 is “absolutely unfair” and 9 “absolutely fair”

of the condition on the decision to convict and on the perceived fairness of the conviction reflect two outcomes of the same cognitive process.⁶⁷

D. Experiment 4 (“Antibiotics”): Forensic Evidence in Civil Litigation

The fourth and final experiment aimed to deal with numerous additional issues that were not fully addressed in the first three experiments. First, it broadened our focus to include civil litigation. Civil litigation is distinct from criminal litigation along several key issues. Of specific relevance to this project is the different evidentiary threshold employed in civil litigation—preponderance of the evidence.⁶⁸ Arguably, the decision threshold might interact with heuristics associated with the treatment of evidence. For example, people might be more open towards basing liability on inferences when only money is on the table and when the plaintiff is only required to establish the case with a probability exceeding 0.5. Thus, an experiment focusing on a civil dispute allows for a greater generalization of our findings.

Second, Experiment 4 extended the inquiry to a case where the question was not whether a wrong had been committed, but rather who had committed it. As the evidence literature has shown, uncertainty regarding the nature of the act and uncertainty regarding the identity of the one who committed it generate distinct effects on the incentives to engage in the regulated behavior.⁶⁹ Hence, fact finders’ behavior with respect to the two types of uncertainty might possibly differ.

Third, as noted above, some jurisdictions set particularly strict limits on drawing an inference upon an inference.⁷⁰ In order to explore whether fact finders share this aversion toward sequential inferences, the present experiment compared not only direct and indirect evidence, but also single and double inference. Furthermore, the comparison between single and double inference allowed us to extend our

($t(115)=3.005$, $p=.003$). Across the two conditions, probability estimates were significantly correlated with fairness judgment ($r=.279$, $p<.01$). The effect of condition (type of evidence) on fairness judgment was not mediated by probability estimates. A regression analysis predicting fairness judgment by condition (direct evidence vs. inference) and subjectively estimated probability yielded significant effects with respect to both probability and condition ($F(114)=7.829$, $p<.01$ and $F(114)=8.508$, $p<.01$ for the effects of the condition and the subjective probability, respectively). As regards the relationship between the conviction decision and the judged fairness, predicting conviction decision by a logistic regression analysis that includes both condition (direct evidence vs. inference) and fairness judgment as independent variables renders the effect of condition nonsignificant ($B=1.278$, $Wald=30.291$, $p<.001$ and $B=.719$, $Wald=1.090$, $p=.297$ for the effects of fairness and condition, respectively). Thus, the difference between the two conditions with respect to conviction rate is no longer significant when we control for judged fairness.

67. See *infra* notes 97–99 and accompanying text.

68. See generally MCCORMICK, *supra* note 6, at 568–69; Eyal Zamir & Ilana Ritov, *Loss Aversion, Omission Bias, and the Burden of Proof in Civil Litigation*, 41 J. LEGAL STUD. 165, 171–82 (2012) (experimentally examining the actual standard of proof in civil litigation).

69. See Henrik Lando, *Does Wrongful Conviction Lower Deterrence?*, 35 J. LEGAL STUD. 327, 329–30 (2006) (distinguishing between mistakes about the act and mistakes about identity).

70. See *supra* notes 13–14 and accompanying text.

understanding of the role of complexity with respect to fact finders' attitude towards circumstantial evidence. Arguably, since a double inference scenario is more complex than a single inference scenario, complexity would suggest a lower willingness to convict based on a double inference. If, however, fact finders' main concern is the type of evidence (direct or circumstantial), then one would not necessarily expect a significant difference between the willingness to convict based on a single or double inference.

Finally, Experiment 4 allowed us to broaden our pool of subjects. While the participants in the previous experiments were first-year law students (or students who completed their first year of legal studies), participants in the present experiment were advanced law students, including master's candidates (many of whom are practicing lawyers). Extending the group of subjects to more diverse populations helps alleviate some of the concerns regarding the external validity of experimental findings.

We used a between-subject design, employing three versions of a vignette describing a civil action, followed by three questions. Responders were asked to imagine that they were serving as a judge in a suit for damages filed by a small dairy against a dairy farmer. The farmer sells the dairy the milk he produces. According to the contract, the farmer should make sure that there are no antibiotics residues in the milk, because such residues obstruct the production of various products. The milk is delivered to the dairy by a tank truck that transports the milk of two farmers. Since the milk of the two farmers is mixed in the tank, a sample is taken from each farmer's milk before pumping it into the tanker, and the samples are delivered to a laboratory, where they will be examined should a problem arise. The dairy claims that the farmer provided milk containing antibiotics residues and, as a consequence, yogurt production failed and the dairy had to discard all of the raw materials it had used in this process. As a result, the dairy suffered a loss of 30,000 NIS.

In the direct-evidence condition, when the yogurt-production process failed, the milk samples of the two farmers were examined. This examination revealed that there were antibiotics residues in the defendant's milk and no residues in the other farmer's milk. The probability that the results of the laboratory examinations are correct (that is, that there were antibiotics residues in the defendant's milk but not in the other farmer's milk) is 85%. It is undisputed that the source of the antibiotics residues could only be milk from one of the farmers.⁷¹

In the single-inference condition, when the yogurt-production process failed, the milk samples from the two farmers were about to be examined (here too, it was undisputed that the source of the antibiotics residues could only be milk from one of them), but it turned out that the sample of the defendant's milk had been lost in the laboratory and, therefore, it was only possible to examine the other farmer's sample. This examination showed no antibiotics residues. The probability that the

71. Arguably, both litigants' acknowledgement that the production failure was caused by their milk may be a product of an inferential process in which other conceivable causes were considered and eliminated. However, in both actual litigation and laboratory experiments, when a fact is "undisputed," fact finders and responders need not, and most likely do not, question the basis of the agreed-upon facts.

results of the laboratory examination are correct (that is, that there were no antibiotics residues in the other farmer's milk) is 85%.

In the double-inference condition, when the yogurt-production process failed, the first possibility examined was that the failure was due to other substances added to the milk during the process. This examination revealed that none of the other substances were tainted. When the milk samples from the two farmers were then about to be examined (here again, it was undisputed that the source of the antibiotics residues could only be milk from one of the farmers), it turned out that the sample of the defendant's milk had been lost in the laboratory and, therefore, it was only possible to examine the other farmer's sample. This examination showed no antibiotics residues. The probability that the results of the laboratory examinations are correct (that is, that the production failure was not caused by other substances or by antibiotics residues in the other farmer's milk) is 85%. Based on the test results of the other substances, the dairy claims that one should infer that the milk caused the failure of the yogurt-production process, and, based on the results relating to the other farmer's sample, one should infer that the defendant was the one responsible.

Given the somewhat greater complexity of the design of this experiment, special attention was given to the wording of the probability of error so as to assure that the conditions did not differ along this dimension. This was achieved by describing the reliability of the laboratory tests in a way that very clearly referred to the *entire* testing process.⁷²

After reading the scenario, responders were asked (1) whether they would accept the claim, (2) what their assessment of the probability that there were antibiotics residues in the defendant's milk was, and (3) how they would rate the fairness of accepting the claim. The three questionnaires were posted on the same website, and participants were recruited in the same way as in Experiment 1. To encourage participation, four participants were randomly selected to win a prize of 150 NIS each. Responders were randomly assigned to one of the three conditions. A total of 193 advanced LLB and LLM students responded to the questionnaire (100 women, 93 men). They ranged in age from 19 to 58, with a mean of 26.

We began analyzing the results by comparing the double-inference group with the single-inference group. The rate of accepting claims—37% in the double-inference condition and 44% in the single-inference condition—did not differ significantly between the two groups.⁷³ Neither did the subjective estimates of liability. While the estimated probability of the existence of antibiotics residues in the defendant's milk was 65.2% in the double-inference group, it was 64.3% in the

72. The three descriptions were as follows. Direct Evidence: "The probability that the results of the laboratory examinations are correct (that is, that there were antibiotics residues in the defendant's milk and no such remains in the other farmer's milk) is 85%"; Single Inference: "The probability that the results of the laboratory examinations are correct (that is, that there were no antibiotics residues in the other farmer's milk) is 85%"; Double Inference: "The probability that the results of the laboratory examinations are correct (that is, that the failure of production was not caused by the other substances or by antibiotics residues in the other farmer's milk) is 85%."

73. $p=0.4$.

single-inference group.⁷⁴ Given the lack of difference between the single- and double-inference conditions, we collapsed the data across the two inference conditions and compared the direct-evidence results to those of the combined-inference conditions.

The rate of accepting the claim differed significantly between the direct-evidence and the inference groups. Whereas 81.7% of the participants in the direct-evidence condition accepted the claim, only 40.2% in the inference conditions did. This difference is highly statistically significant.⁷⁵ The estimated probability of the existence of antibiotics residues in the defendant's milk differed significantly between the two conditions as well. While the mean estimate in the direct-evidence condition was 81.7%, the equivalent figure with respect to the inference conditions was only 64.8%.⁷⁶ Not surprisingly, across all conditions, participants who indicated that they would accept the claim evaluated the probability of the presence of antibiotics residues in the defendant's milk higher than participants who did not accept the claim (80.6% and 69.1%, respectively).⁷⁷

Nonetheless, although the subjective probability estimates are closely related to the decision to accept the claim, the difference between the conditions with respect to the tendency to accept the claim was not mediated by the difference in probability estimates. A logistic regression predicting the decision to accept the claim from probability estimate and condition (direct evidence vs. inference, coded as in Experiment 1) yielded significant effects with respect to both factors.⁷⁸ In other words, the tendency to assign legal liability remained significantly lower in the circumstantial evidence condition even when controlling for fact finders' subjective probability estimates.⁷⁹ As in Experiment 3, the effect of condition, direct evidence versus inference, was possibly mediated by the effect of this factor on judged fairness.⁸⁰

74. $p=0.8$.

75. Chi-Square=31.328, $p<0.001$.

76. $t(191)=6.663$, $p<0.001$.

77. $t(191)=8.119$, $p<0.001$.

78. $B=.054$, $Wald=25.666$, and $B=1.296$, $Wald=10.976$, for probability estimate and condition, respectively, with $p<.01$ for both factors.

79. These results differ from the results of the second experiment reported by Arkes et al., *supra* note 44, at 282–84. In that experiment, Arkes and his coauthors used vignettes similar to the tire-tracks scenario of Wells's experiment, *supra* note 36, at 743, yet while in one condition the witness found that the tire tracks matched the blue bus company vehicles, in the other the witness found that the tracks did not match the grey bus company vehicles. In both conditions, the technique used was correct 80% of the time. Responders' willingness to impose liability on the blue bus company was more than twice as great in the first condition, and this difference was mediated by their different probability assessments. In other experiments Arkes and his coauthors describe, the verdict was affected by various factors *without* being mediated by the assessed probability. Arkes et al., *supra* note 44, at 280–82 and 285–88.

80. Predicting a decision to accept the claim by a logistic regression analysis that includes both condition (direct evidence vs. inference) and fairness judgment as independent variables renders the effect of condition nonsignificant ($B=1.509$, $Wald=54.372$, $p<.001$ and $B=.445$, $Wald=.551$, $p=.458$ for the effects of fairness and condition, respectively).

III. DISCUSSION AND POLICY IMPLICATIONS

This Part examines the contribution of our findings to the legal literature. We first outline our main results and discuss them in light of existing theories regarding the prevalent dissimilar treatment of direct and circumstantial evidence. That done, we turn to review the potential normative implications that can be drawn from them. Finally, we highlight some of the limitations of the study and sketch out paths for future research.

A. General Discussion

All four experiments demonstrated an anti-inference bias: the respondents were far more reluctant to impose liability based solely on circumstantial evidence than on direct evidence. This reluctance was manifest across different scenarios; was true for both criminal and civil liability; was common to technological, scientific, and eyewitness evidence; and was evident both when the disputed question was whether a wrong had been committed and when the question was who had committed the wrong. In all four experiments, the anti-inference bias characterized case-specific circumstantial evidence rather than merely statistical evidence. We found a strong disinclination to base liability on circumstantial evidence in both between- and within-subject experimental designs; in both ex post judicial decision making and ex ante policy making; and both when responders made the decision themselves and when they specified how another person should decide. This was true of students at the beginning of their first year of law school and of advanced LLB and LLM students. Furthermore, while the type of evidence, direct or circumstantial, sometimes affected the subjectively assessed probability of the defendant's fault, in none of the experiments did it mediate the effect of the type of evidence on the decision whether to impose liability.

As the following analysis demonstrates, our findings extend the existing literature on circumstantial evidence. While the explanations for the differential treatment accorded to circumstantial evidence put forward by this body of literature might be valid, they cannot account for the full scope of this phenomenon. The experiments suggest that there is another aspect of human decision making that is driving fact finders' behavior.

The first body of explanations highlighted by the existing literature focuses on objective probabilities or on fact finders' subjective probability assessments.⁸¹ Our findings, however, suggest that these factors cannot fully account for people's reluctance to impose legal liability based on indirect evidence. In Experiment 1 (Highway), Experiment 3 (Bus), and Experiment 4 (Antibiotics), the objective probability that the defendant had committed the act attributed to him was identical in the direct-evidence and the inference conditions, yet the responders' willingness to impose legal liability differed markedly. In the same vein, despite the identical accurateness of the two enforcement systems in Experiment 2 (Enforcement), the responders strongly preferred the direct-evidence system over the inference one. The assumption that circumstantial evidence is inherently less conclusive than

81. See *supra* notes 23–25 and accompanying text.

direct evidence thus cannot explain the prevalent disinclination to impose liability based on such evidence.

As for subjective probability assessments, despite the fact that there was no significant difference between responders' mean probability assessments in the two conditions in Bus, about 65% of the responders found the driver guilty of violating the terms of the permit in the direct-evidence condition, but only 38% did so in the inference condition. In Antibiotics, while there were significant differences in the subjective probability assessments between the two conditions, these differences did not mediate the difference in the responders' willingness to accept the claim under the two conditions, as this difference remained significant even when we controlled for the differences in probability assessments. In this regard, our findings are in line with previous results;⁸² yet, they generalize the previous results and expose a wide-ranging bias against circumstantial evidence that is not limited to the context of statistical and forensic evidence. In all of our experiments, conditions differed with regard to the directness of the evidence, but not with regard to whether the evidence was statistical, scientific, or eyewitness. Our findings also extend the previous experimental findings in covering criminal cases (Highway, Enforcement, and Bus) in addition to civil ones (Antibiotics); an *ex ante* choice between enforcement systems (Enforcement) in addition to *ex post* judicial decisions (Highway, Bus, and Antibiotics); and within-subject (Enforcement) in addition to between-subject design (Highway, Bus, and Antibiotics).

The incompatibility of our findings with the subjective probability explanation indicates that explanations building on the greater complexity of circumstantial evidence cannot fully account for the responders' disinclination to rely on this type of evidence. Had the fact finders been reluctant to impose liability based on circumstantial evidence because of their difficulty in understanding and drawing conclusions from such evidence, one would expect this difficulty to affect, first and foremost, the judged probability of the defendant's fault, and that we would find a correlation between the subjective probability assessment and the responders' willingness to impose liability. Moreover, while in Antibiotics the circumstantial evidence (particularly in the double-inference condition) was indeed more complex than in the direct-evidence condition, this can hardly be said of Highway, Enforcement, and Bus. Finally, in Antibiotics there was a considerable difference between the single-inference and double-inference conditions in terms of their complexity, but no significant difference in terms of the responders' willingness to impose liability. Thus, complexity and clarity of the evidence cannot fully account for the general phenomenon we identified in our experiments.

Another explanation for the different treatment given to circumstantial evidence builds on the abstract and pallid nature of such evidence.⁸³ However, the evidence presented in all four experiments in both conditions was of the exact same character (e.g., eyewitness testimony or scientific/technological) and did not involve naked statistical evidence. This would indicate that the explanations based on the relative representational versus abstract, narrative versus rhetorical, unconditional versus

82. See Arkes et al., *supra* note 44, at 280–90; Niedermeier et al., *supra* note 38, 534–41; Wells, *supra* note 36, at 741–48.

83. See *supra* notes 26–29 and accompanying text.

probabilistic, vivid versus pallid, concrete versus general, and emotional versus unexciting nature of the evidence are incomplete. For instance, in Highway and Enforcement, the evidence produced by the one-camera and two-camera systems was equally abstract, probabilistic, pallid, concrete, and unexciting.

Similar observations apply to the case-specificity argument, according to which naked statistical evidence is insufficient for imposing liability, even when the probability is very high, because the weight or resiliency of the evidence is equally important. According to this argument, if the assessed probability rests on too little information or on general, non-case-specific evidence, it should not be relied upon in order to determine liability.⁸⁴ While it is easy to equate naked statistical evidence and circumstantial evidence, the two represent distinct issues. None of our experiments included naked statistical evidence. Rather, they all involved case-specific evidence that was either direct or circumstantial. It is unclear why, for example, in Highway one would consider the results of the two-camera system as having less weight or as less resilient than the results of the one-camera system.

Our findings also suggest that the reluctance to impose liability based on circumstantial evidence is not necessarily a matter of responsibility taking.⁸⁵ According to the responsibility-taking theory, when relying on direct evidence, the fact finder feels less accountable for erroneous judgments because it is not her inference, but rather the witness's fallibility or dishonesty that brought about the regrettable judicial error. As a result, the fact finder is less hesitant to impose liability based on direct evidence.

When viewing the decisions subjects made in all experiments through the prism of responsibility taking, it is clear that there was no real difference between the two conditions used in the Highway, Bus, and Antibiotics scenarios in this regard. Certainly no issue of responsibility taking was involved in the Enforcement experiment, where responders were asked to choose between two speed-enforcement systems *ex ante*—in the abstract—rather than to impose liability on any particular person. Furthermore, in Bus, responders were not asked to make the decision themselves, but rather to indicate how *another person*—the judge—should decide. This did not, however, counteract the responders' reluctance to rely on circumstantial circumstances.⁸⁶

The experiments show that the fact-to-evidence theory cannot fully account for fact finders' behavior either.⁸⁷ The fact-to-evidence theory posits that people are far more inclined to rely on evidence when the decision of whether to impose liability bears on the reliability of the evidence. Our findings, however, do not support this conjecture.⁸⁸ For instance, in both conditions of the Bus experiment, convicting the

84. *See supra* notes 30–31 and accompanying text.

85. *See supra* notes 32–34 and accompanying text.

86. Our vignettes also varied in terms of the identity of the person who drew the inference. In Highway and Enforcement, it was “possible to infer” from the indirect evidence; in Bus, “the policeman inferred”; and in Antibiotics the plaintiff urged the court (that is, the responder) to infer. Arguably, responders should have experienced greater personal responsibility in Antibiotics and less in Bus. In fact, the disinclination to impose inference-based liability was common to all four scenarios.

87. *See supra* notes 35–36 and accompanying text.

88. In this regard our findings corroborate the findings of Niedermeier, Kerr, and Messé.

driver implied that the policeman's testimony was very reliable, and acquitting the driver implied that it was insufficiently reliable; yet there was a huge difference between the two conditions in the willingness to convict.⁸⁹

We then turn to the ease-of-simulation theory.⁹⁰ As originally conceived by Daniel Kahneman and Amos Tversky⁹¹ and subsequently studied in various contexts,⁹² the simulation heuristic pertains to how likely one thinks it is that an outcome will occur or an event has occurred (or could have occurred). According to the theory, the easier it is to imagine a scenario (and the harder it is to imagine a competing one), the more probable it appears to be.⁹³ Several scholars have tied this pattern of thinking to the way in which fact finders treat circumstantial evidence.⁹⁴ The core argument is that, by its very nature, circumstantial evidence makes it easier to imagine alternative scenarios in which the defendant is not liable. As a result, fact finders will systematically under-assess the probability that the defendant is liable when asked to base liability on circumstantial evidence.

The ease-of-simulation heuristic cannot explain fact finders' disinclination to impose liability based on circumstantial evidence in the experiments reported in this Article. As both Bus and Antibiotics suggest, this disinclination is not based on a cognitive process that is tied to participants' judgments of probabilities. In the former, no difference was found between the direct and the circumstantial groups with respect to probability assessments; in the latter, the disinclination remained significant even when controlling for the differences in subjective probability. Explaining the reluctance to base liability on purely circumstantial evidence by the

See Niedermeier et al., *supra* note 38, at 536–39.

89. In Enforcement, it could hardly be argued that choosing one system necessarily implied that it was more accurate than the alternative. In that experiment, 38 of 49 responders preferred one of the systems (one-camera/direct system or two-camera/inference system) to the other. Of those 38, 19 (50%) would stick with their preferred choice even if its accuracy were lower than that of the alternative.

90. See *supra* notes 35–41 and accompanying text.

91. Kahneman & Tversky, *supra* note 37.

92. See, e.g., Ariel S. Levi & John B. Pryor, *Use of the Availability Heuristic in Probability Estimates of Future Events: The Effects of Imagining Outcomes Versus Imagining Reasons*, 40 *ORG. BEHAV. & HUM. DECISION PROCESSES* 219, 222–29 (1987) (examining the factors affecting people's prediction of who will win the elections); David Raune, Andrew MacLeod & Emily A. Holmes, *The Simulation Heuristic and Visual Imagery in Pessimism for Future Negative Events in Anxiety*, 12 *CLINICAL PSYCHOL. & PSYCHOTHERAPY* 313, 315–21 (2005) (studying the assessed probability of future negative events by patients suffering from clinical anxiety).

93. The ease of simulation may depend on the ease of simulating other similar events, Dale T. Miller, William Turnbull & Cathy McFarland, *When a Coincidence Is Suspicious: The Role of Mental Simulation*, 57 *J. PERSONALITY & SOC. PSYCHOL.* 581, 583–86 (1989), and counterfactuals, Edward R. Hirt & Keith D. Markman, *Multiple Explanation: A Consider-an-Alternative Strategy for Debiasing Judgments*, 69 *J. PERSONALITY & SOC. PSYCHOL.* 1069, 1072–83 (1995), as well as on the perceived normality of different events, Daniel Kahneman & Dale T. Miller, *Norm Theory: Comparing Reality to Its Alternatives*, 93 *PSYCHOL. REV.* 136, 137 (1986); Miller et al., *supra* note 93. The ease of simulation is to some extent manipulable. See, e.g., Niedermeier et al., *supra* note 38, at 539–41.

94. See, e.g., Heller, *supra* note 1; Niedermeier et al., *supra* note 38; Sykes & Johnson, *supra* note 40.

ease-of-simulation heuristic thus does not fall in line with the conventional meaning of this heuristic that is directly tied to an estimate of the likelihood of an event.

Moreover, to a large degree our vignettes were designed to eliminate alternative explanations and thus rule out the influence of this heuristic. For instance, in the inference condition in Highway, given the distance between the two cameras and the time that elapsed between the documenting of the car by each camera, there is simply no way the driver did not violate the speed limit. True, one could imagine an error in the system, but the phrasing of both conditions was identical in this respect: “The probability of an error in the camera system is 2%.” Similarly, in the inference condition of Bus, the only explanation for the fact that four tourists were left behind after the two minibuses that arrived empty were filled up and drove away is that there were fifty-four passengers on the tourist bus.⁹⁵ Any alternative explanation was ruled out, as the questionnaire explicitly stated that fifty people were driven away on the two minibuses.⁹⁶

Last we consider the fairness explanation arguably supported by our findings. In Highway, Bus, and Antibiotics there was a statistically significant positive correlation between subjects’ decisions to impose liability and their judged fairness of doing so. However, it is unclear how this correlation should be interpreted. One possibility is that imposition of liability is unfair because it is prone to error due to either the inherent inconclusiveness of circumstantial evidence or its lesser weight and resiliency. However, this does not seem to be the notion of fairness envisioned by the responders, as in none of the experiments was the imposition of liability mediated by the assessed probability that the defendant actually committed the act, and in all experiments the evidence was case specific.⁹⁷

95. The lack of a statistically significant difference between the subjective probability assessments in the two conditions suggests that subjects did not view an alternative scenario as plausible.

96. Interestingly, despite its centrality in the literature, the ease-of-simulation theory fails to provide a complete explanation for the different treatment of direct and circumstantial evidence even in Niedermeier and his coauthors’ own study. Niedermeier et al., *supra* note 38. Their study directly tested the mediating effect of the ease-of-simulation variable on verdicts in two of its experiments. While in both experiments the ease of simulation affected the verdicts, in neither did it fully mediate the effect of type of evidence. *Id.* at 536–41. Although controlling for ease of simulation considerably reduced the proportion of variance in the verdicts, there remained a statistically significant difference between the verdicts in the different conditions. *Id.* at 538, 541. Thus, even according to Niedermeier and his coauthors, the ease of simulation provides an incomplete explanation for factfinders’ behavior. A more elaborate examination of the ease-of-simulation explanation was recently conducted by Arkes and his coauthors. Arkes et al., *supra* note 44. They found that sometimes this factor (which they dub “imaginability”) possibly mediates the verdict, yet sometimes it is unrelated to the verdict or even related in a manner completely contrary to the ease-of-simulation hypothesis. *Id.* at 282–86.

97. It may also be argued that in Highway, the speed-camera system was perceived as fairer because it is more forgiving: It does not penalize drivers who violate the speed limit in the parts of the highway not covered by a speed camera, that is, most of the road. However, the two-camera system is more forgiving in the sense that it only captures the driver’s *average* speed in the pertinent section of the road, which is typically lower than the highest

A different notion of fairness might have to do with people's expectations and the "rules of the game." Perhaps people believe that they should only be penalized if they are caught in the act. Imposing an inference-based liability is thus seen as unfair, regardless of the reliability of the circumstantial evidence. While such a notion of fairness possibly plays a role in some contexts, it can hardly serve as a general explanation for the reluctance to impose inference-based liability. First, it is unclear that, in deciding whether to impose criminal or civil liability, people consider the wrongdoer's expectations as to the type of evidence that might be used against her. Second, when one's behavior harms another person, as in *Antibiotics* and in Wells's blue bus scenario,⁹⁸ not compensating the injured party because the evidence supporting her claim is not the type expected by the defendant seems considerably more unfair than using such evidence. Finally, the present notion of fairness does not account for the disinclination to rely on circumstantial evidence even when the pertinent evidence, such as forensic evidence found at the scene of the crime, has long been used by the police and is now part of the "rules of the game."⁹⁹

At the end of the day, it may well be that the responders in our experiments understood "fair" as roughly equivalent to "appropriate" or "acceptable." The decision to impose criminal or civil liability and the characterization of the imposition of liability as fair echoed one another, possibly reflecting the same cognitive process. Thus, the answer to the fairness question was possibly a post hoc rationalization of responders' answers to the verdict question.

To sum up, our findings show that the reluctance of fact finders to impose liability based solely on indirect, circumstantial evidence is stronger and further reaching than demonstrated in previous studies. None of the explanations proposed thus far for this reluctance are completely satisfactory. Our findings reveal an independent, deeply rooted anti-inference heuristic.

Much like other simplifying heuristics, the anti-inference heuristic assists people to make judgments under uncertainty. It functions as a substitute for extensive algorithmic processing and yields judgments that are usually accurate. Ordinarily, when we see something with our own eyes, or when someone tells us that she saw something herself, that event actually happened; this is not necessarily true of conclusions drawn from circumstantial evidence. We tend to trust our power of vision more than we trust our power of deduction. However—like other biases and heuristics—the anti-inference heuristic also gives rise to systematic error; it strongly affects decisions even when the objective and subjective probabilities of the pertinent occurrence are equal according to circumstantial and direct evidence.¹⁰⁰ The anti-inference heuristic may thus be seen as yet another

speed she drove in that section.

98. Wells, *supra* note 36.

99. The rules of the game may change over time. Thus, while speed cameras have come to be regarded as a commonplace in Israel (where the experiments were conducted), two-camera systems of the kind described in *Highway and Enforcement* are not ordinarily used in Israel and can thus be perceived as unfair. Even if this conjecture helps to explain the results of *Highway and Enforcement*, it does not seem relevant to *Bus* or *Antibiotics*.

100. There is an intriguing correspondence between our findings and the findings of experiments studying young children's overly rigid application of a "seeing = knowing" rule according to which facts can only be determined through direct evidence and not through

manifestation of overgeneralization. It reflects an inability of people to perfectly analyze the decision environment in which they function. While preferring to rely on direct evidence might be a prudent strategy in day-to-day life, in some contexts, it might be a problematic decision-rule.

B. Policy Implications

Having described the positive findings of our experiments, we now turn to analyze their potential normative implications.¹⁰¹ First, we explore the implications for the rules governing the structure of trials. In this regard, we examine whether legal systems can adopt strategies that will debias fact finders. We then turn to examine the implications for evidence law and for substantive legal rules. As the discussion will show, legal systems anticipating the anti-inference bias adopt legal presumptions and structure prohibitions that, to a certain degree, circumvent it.

1. Debiasing Fact Finders

The existence of a systematic aversion towards basing liability on circumstantial evidence suggests that like other heuristics, the anti-inference bias might drive judicial decision making astray. Thus, legal systems will be able to improve the quality of fact finding by circumventing the bias through proper regulation. Just as the legal system might try to overcome cognitive phenomena such as the hindsight bias,¹⁰² it might also try to debias fact finders influenced by the anti-inference bias.

A straightforward avenue to attempt counteracting the anti-inference bias is instructing fact finders to afford equal weight to direct and circumstantial evidence. Numerous jurisdictions have taken this path by stressing in their jury instructions that direct and circumstantial evidence are in principle equally reliable.¹⁰³ In Connecticut, for example, the instructions state that: “There is no legal distinction between direct and circumstantial evidence as far as probative value; the law

inference. *See, e.g.*, Ori Friedman, Richard Griffin, Hiram Brownell & Ellen Winner, *Problems with the Seeing = Knowing Rule*, 6 DEVELOPMENTAL SCI. 505, 506–07 (2003) (presenting a critical review of the literature on the matter); Ted Ruffman, *Do Children Understand the Mind by Means of Simulation or a Theory? Evidence from Their Understanding of Inference*, 11 MIND & LANGUAGE 388, 393–400 (1996) (reporting experimental findings documenting the phenomenon).

101. To be sure, in the following analysis we focus our discussion on the possibility that fact finders exhibit excessive caution towards circumstantial evidence, and not excessive confidence towards direct evidence. To the extent the latter option reflects the decision-making pattern, the proposals outlined in this Part should be adjusted accordingly.

102. *See, e.g.*, D. Jordan Lowe & Philip M.J. Reckers, *The Effects of Hindsight Bias on Jurors' Evaluations of Auditor Decisions*, 25 DECISION SCI. 401, 408–16 (1994) (presenting experimental findings on successful debiasing in the hindsight context); David B. Wexler & Robert F. Schopp, *How and When to Correct for Juror Hindsight Bias in Mental Health Malpractice Litigation: Some Preliminary Observations*, 7 BEHAV. SCI. & L. 485, 489–92 (1989) (exploring debiasing strategies in the context of the hindsight bias).

103. *See, e.g.*, Jacquelyn L. Bain, *A Proposed Definition of Reasonable Doubt and the Demise of the Circumstantial Evidence Charge Following Hankins v. State*, 15 ST. MARY'S L.J. 353, 362 (1984) (noting the equal weight now afforded to circumstantial evidence in different jurisdictions).

permits you to give equal weight to both, but it is for you to decide how much weight to give to any particular evidence.”¹⁰⁴ Arguably, legislatures specifying such instructions both acknowledge the existence of the anti-inference bias, and attempt to deal with it by offering jurors proper guidance.

The effectiveness of such instructions, however, is not self-evident. The research on jury instructions in other areas suggests that instructions are often ineffective since juries tend to disregard them.¹⁰⁵ In the context of the right to remain silent, for example, studies have documented the tendency of jurors to hold the exercise of this right against defendants despite explicit instructions not to do so. Thus, after conducting a series of stylized experiments on the issue, David Shaffer concluded that “this bias against defendants who invoked the Fifth Amendment was apparent even though the judge had affirmed the defendant’s right to so plead and had instructed jurors that they were to draw no inferences about the defendant’s innocence or guilt from his use of this constitutional privilege.”¹⁰⁶

Furthermore, in some cases the strategy of jury instructions might even backfire and do more harm than good.¹⁰⁷ In the context of information about previous convictions, for instance, Kerri Pickel demonstrated that people are more likely to convict after being exposed to information regarding past convictions if they are instructed to disregard that information.¹⁰⁸ Similarly, once the circumstantial nature of evidence is made salient by the instructions, jurors might be drawn to discount their probative value.

A second path policy makers might choose to take in order to deal with the anti-inference bias is to bring it to the attention of fact finders. According to this line of thought, courts should allow for expert witness testimony that will explain to fact finders the precise pitfalls associated with the anti-inference bias.¹⁰⁹ Arguably, a detailed scientific explanation on the bias might prove more effective than standard jury instructions that are easy to ignore. From a legal perspective, allowing for expert testimony regarding the anti-inference bias should not face insurmountable difficulties. Rule 702 of the Federal Rules of Evidence grants judges broad discretion over the admissibility of expert witness testimony.¹¹⁰ That said, one

104. CONN. CRIM. JURY INSTRUCTIONS § 2.4-1 (2007) (modified 2008).

105. See Joel D. Lieberman & Jamie Arndt, *Understanding the Limits of Limiting Instructions: Social Psychological Explanations for the Failures of Instructions to Disregard Pretrial Publicity and Other Inadmissible Evidence*, 6 PSYCHOL. PUB. POL’Y & L. 677, 703 (2000) (stating that “the majority of extant empirical research indicates that jurors do not adhere to limiting instructions”).

106. David R. Shaffer, *The Defendant’s Testimony*, in THE PSYCHOLOGY OF EVIDENCE AND TRIAL PROCEDURE 124, 143 (Saul M. Kassir & Lawrence S. Wrightsman eds., 1985).

107. See Lieberman & Arndt, *supra* note 105, at 689–91 (reviewing the evidence on the “backfire” effect caused by limiting instructions).

108. Kerri L. Pickel, *Inducing Jurors to Disregard Inadmissible Evidence: A Legal Explanation Does Not Help*, 19 LAW & HUM. BEHAV. 407, 411–15 (1995).

109. For a similar proposal in the context of the hindsight bias, see Wexler & Schopp, *supra* note 102, at 490–92.

110. See FED. R. EVID. 702 (establishing a broad standard that enables courts to admit evidence that can help the trier of fact “understand the evidence or to determine a fact in issue”).

should acknowledge the limitation of this solution. Expert evidence is costly,¹¹¹ and it is unclear whether litigating parties will find it beneficial to invest the necessary resources in producing it. When the stakes are relatively low, as they so often are, an expert witness will probably not justify the costs associated with summoning her to court.

Finally, while the issue of debiasing raises numerous practical questions relating to the feasibility of such policies, it also raises a more principled dilemma. If the anti-inference bias reflects deeply held epistemological and moral intuitions, circumventing it might adversely affect the desirable correspondence between peoples' prevalent perceptions and the outcomes of adjudication. As Jeffrey Rachlinski concluded, a judgment produced by a debiased trier of fact "might seem less fair than an uncorrected, biased judgment."¹¹² Our findings substantiate this concern. In Highway, Bus, and Antibiotics, we asked the responders not only to decide (or indicate how a judge should decide) the case, but also to assess the fairness of imposing liability. In all three experiments, there was a statistically significant positive correlation between the answers to the two questions: The fairer the responders judged the imposition of liability, the likelier they were to convict the defendant or accept the claim. Put differently, people's reluctance to impose liability based on indirect evidence correlates with their belief that imposing such liability is less "fair."¹¹³

Counteracting the anti-inference bias might therefore adversely affect the legitimacy and effectiveness of the legal system. A significant body of work has demonstrated the importance of the perceived fairness of the criminal justice system for its effectiveness.¹¹⁴ More recently, scholars have extended this argument and suggested that perceived fairness might play a key role in fostering cooperation in civil contexts as well.¹¹⁵ Deciphering the precise connections between the anti-inference bias, circumventing it, and perceptions of fairness, lies beyond the scope of this Article. Nonetheless, any legal treatment of the issue will have to take this issue into account.

111. See, e.g., Monique Michal, Comment, *After West Virginia: The Fate of Expert Witness Fee Shifting in Patent Litigation*, 59 U. CHI. L. REV. 1591, 1591 (1992) (reviewing the high costs associated with expert witnesses).

112. Jeffrey J. Rachlinski, *A Positive Psychological Theory of Judging in Hindsight*, 65 U. CHI. L. REV. 571, 601 (1998).

113. See *supra* notes 52–53, 66–67, 80, 97–98 and accompanying text.

114. See, e.g., Josh Bowers & Paul H. Robinson, *Perceptions of Fairness and Justice: The Shared Aims and Occasional Conflicts of Legitimacy and Moral Credibility*, 47 WAKE FOREST L. REV. 211, 212–13 (2012) (pointing to the "significant crime-control advantages for a system that enjoys perceptions of both moral credibility and legitimacy"); Paul H. Robinson & John M. Darley, *Intuitions of Justice: Implications for Criminal Law and Justice Policy*, 81 S. CAL. L. REV. 1, 18–32 (2007) (reviewing the undesirable implications of policies that distribute punishment in a way that does not accord with shared intuitions of justice); Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453, 454 (1997) (advocating a criminal law that reflects "the community's perceptions of just desert").

115. See, e.g., Yuval Feldman & Doron Teichman, *Are All Contractual Obligations Created Equal?*, 100 GEO. L.J. 5, 36 (2011) (noting the connection between perceived fairness and contractual compliance).

2. Rules of Evidence

Aware of the challenges associated with direct debiasing, legal policy makers may resort to indirect measures in order to deal with the anti-inference bias. The obvious example is the introduction of legal presumptions. The law often resorts to presumptions in order to assign liability.¹¹⁶ A presumption allows “the existence of an ultimate fact, or presumed fact, to be taken as established where other predicate or basic facts have been established.”¹¹⁷ Thus, for example, defendants who cannot present a reasonable explanation for the fact that they are in possession of recently stolen goods might be convicted of stealing those goods based on a presumption deducing theft from possession.¹¹⁸ Similarly, drugs that are found in the defendant’s home might be assumed to be in his possession based on a presumption connecting the two.¹¹⁹ Interestingly, the use of such presumptions is not limited to jurisdictions that delegate fact finding to nonprofessional jurors.¹²⁰

Despite their prevalence, the existence of legal presumptions is somewhat puzzling. As courts often note, presumptions rest on the existence of a “common-sense” connection between the fact that was established and the fact that is deduced from it.¹²¹ Actually, such a connection is a precondition for the constitutionality of the presumption.¹²² On its face, there is no need to instruct fact finders to use common sense when determining facts, since that is what they are expected to do whether instructed or not. As one judge noted, she would be “hard-pressed” if asked by jurors for clarifications regarding how they should implement a presumption.¹²³

The anti-inference bias offers an explanation for the pervasiveness of legal presumptions.¹²⁴ Given the ability of presumptions “‘to encourage’ the jury to find

116. In the following discussion we focus on legal presumptions in the context of criminal law, though this phenomenon clearly goes well beyond this body of law. For example, in family law presumptions play an important role with respect to key issues. See Brie S. Rogers, Note, *The Presumption of Paternity in Child Support Cases: A Triumph of Law over Biology*, 70 U. CIN. L. REV. 1151, 1151–52 (2002) (noting that the presumption of paternity is deeply rooted in American legal tradition).

117. Theodore A. Gottfried & Peter G. Baroni, *Presumptions, Inferences and Strict Liability in Illinois Criminal Law: Preempting the Presumption of Innocence?*, 41 J. MARSHALL L. REV. 715, 715 (2008).

118. See, e.g., Lynn McLain, “Other Acts” Evidence: Recent Decisions by the Court of Appeals Undermine the Efficacy of Maryland Rule 5-404(b), 31 U. BALT. L.F. 5, 9 (2000) (noting that this is a well-established presumption under Maryland law).

119. See, e.g., *United States v. Hill*, 142 F.3d 305, 312 (6th Cir. 1998).

120. See, e.g., CA 9282/00 Yarchi v. State of Israel PD 55(5) 759 [2001] (Isr.) (Israeli Supreme Court ruling that presence at the scene of the crime creates a presumption that the defendant was an accomplice to the crime).

121. *Proof Issues*, 40 GEO. L.J. ANN. REV. CRIM. PROC. 696, 706 (2011) (reviewing the case law).

122. *Id.*

123. *People v. Thomas*, 407 N.Y.S.2d 812, 814 (N.Y. Crim. Ct. 1978).

124. We do not argue that the anti-inference bias is the sole explanation for the existence of legal presumptions. For other arguments see, for example, J. Harvie Wilkinson III, *Toward A*

fact *B* . . . once fact *A* . . . is proven,”¹²⁵ one could view them as a mild form of regulation that counteracts the effect of the bias. By adopting presumptions, the law pushes fact finders to make common-sense deductions they might otherwise be reluctant to make. Thus, presumptions help steer fact finders towards more accurate decisions.

3. Substantive Legal Rules

Thus far our analysis has focused on the implications of the anti-inference bias on adjudication. This mode of analysis is in line with traditional legal scholarship that envisions a division of labor between substantive legal rules and the rules of evidence and procedure.¹²⁶ Whereas substantive legal rules aim to promote the general goals of the legal system such as fairness, just desert, and efficiency, the rules of evidence and procedure focus on advancing the goals of the judicial process, such as enhancement of accuracy of fact finding and desirable apportionment of the risk of error.¹²⁷

Recent research, however, has demonstrated that this dichotomous view of the legal system is inaccurate. On one hand, the rules governing the legal process are, or at least should be, attuned to the primary goals of the legal system.¹²⁸ For example, the rules of evidence might mediate between deterrence and retribution as the guiding principle of criminal law.¹²⁹ On the other hand, substantive rules might be used to reduce the risk of judicial error. For instance, sanctions can be calibrated such that they reduce the probability of wrongful convictions.¹³⁰

In the context of the anti-inference bias, policy makers aware of its existence might choose to deal with it by structuring prohibitions in a way that will overcome fact finders’ reluctance to deduce liability. One strategy in this regard is to broaden the scope of prohibitions, such that they not only encompass the primary behavior that the policy maker wishes to regulate, but also related behaviors from which one could deduce that the primary behavior took place. By adopting such a legislative strategy, the state no longer needs to persuade fact finders to infer liability, as the inference was already determined by the legislature.

Jurisprudence of Presumptions, 67 N.Y.U. L. REV. 907 (1992) (arguing that presumptions function as a desirable middle point between rule based regimes and equity based regimes).

125. Ronald J. Allen, *Structuring Jury Decisionmaking in Criminal Cases: A Unified Constitutional Approach to Evidentiary Devices*, 94 HARV. L. REV. 321, 335 (1980).

126. *See, e.g.*, Ehud Guttel & Doron Teichman, *Criminal Sanctions in the Defense of the Innocent*, 110 MICH. L. REV. 597, 609–10 (2012) (highlighting the perceived division of labor within law).

127. *See, e.g.*, STEIN, *supra* note 30, at 1.

128. *See, e.g.*, Louis Kaplow, *Burden of Proof*, 121 YALE L.J. 738, 762–72 (2012) (analyzing, *inter alia*, the effect of burden of proof on primary activities); Chris William Sanchirico, *A Primary-Activity Approach to Proof Burdens*, 37 J. LEGAL STUD. 273, 280–86 (2008) (same).

129. *See generally* Richard A. Bierschbach & Alex Stein, *Mediating Rules in Criminal Law*, 93 VA. L. REV. 1197 (2007).

130. *See* Guttel & Teichman, *supra* note 126, at 607–10.

Criminalizing preparation to commit a crime is a case in point. Preparatory offenses are closely related to the law of criminal attempts.¹³¹ Criminalizing attempts promotes important policy goals such as deterrence and prevention. From a deterrence perspective, it allows punishing offenders even when they fail to complete their plan, and thus cheaply raises the probability of punishment.¹³² From a prevention perspective, it enables the police to intervene prior to the completion of the criminal act and empowers the state to incapacitate individuals who demonstrate a propensity for criminal activity.¹³³ Thus, it is unsurprising that most jurisdictions criminalize attempts in one way or another.¹³⁴

Securing convictions in attempt cases might, however, prove a thorny task since they always involve situations in which at least one of the objective elements of the crime is absent.¹³⁵ According to the Model Penal Code, to commit an attempt, an actor must take “a substantial step in a course of conduct planned to culminate in his commission of the crime.”¹³⁶ Thus, in cases involving incomplete attempts, fact finders are left to deduce missing elements of the criminal plan from circumstantial evidence. If, for example, a suspect is apprehended near a bank with a gun and a demand note, the fact finder will have to infer based on circumstantial evidence what would have happened had the police not intervened. Fact finders subject to the anti-inference bias are expected to exhibit a special aversion towards convicting people of such attempts.¹³⁷

As policy makers become aware of the reluctance to convict in criminal attempt cases, they can try to overcome this reluctance by redesigning criminal prohibitions. Specifically, they could define certain acts of preparation as independent crimes. Once the possession of burglary tools or the conducting of a sexual conversation over the Internet with a minor are defined as crimes,¹³⁸ fact

131. Andrew Ashworth, *Conceptions of Overcriminalization*, 5 OHIO ST. J. CRIM. L. 407, 414 (2008) (noting that “[t]he essence of a preparatory crime is a criminal attempt”).

132. See Steven Shavell, *Deterrence and the Punishment of Attempts*, 19 J. LEGAL STUD. 435, 436–37 (1990).

133. See WAYNE R. LAFAVE, CRIMINAL LAW 581–83 (4th ed. 2003) (analyzing criminal attempts from the perspective of early prevention); Shavell, *supra* note 132, at 458 (analyzing criminal attempts from the perspective of incapacitation).

134. R.A. DUFF, CRIMINAL ATTEMPTS 1 (1996).

135. See Arnold N. Enker, *Impossibility in Criminal Attempts—Legality and the Legal Process*, 53 MINN. L. REV. 665, 670 (1969).

136. MODEL PENAL CODE § 5.01(1)(c) (1962).

137. For examples of cases demonstrating the difficulty to assign liability in such situations, see *United States v. Harper*, 33 F.3d 1143, 1147–48 (9th Cir. 1994) (defendants sitting in a car outside of a bank at 10:00 PM equipped with two handguns, ammunition, a roll of duct tape, a stun gun, and a pair of latex surgical gloves after tampering with the bank’s ATM in order to draw the service personnel to the bank, did not take a substantial step toward robbing the bank); *United States v. Still*, 850 F.2d 607, 608 (9th Cir. 1988) (defendant sitting in a car near a bank wearing a blond wig, equipped with a fake bomb, a pouch with a demand note taped to it, a police scanner, and a notebook containing drafts of demand notes, and who later confessed that the police caught him “five minutes before [he] was going to rob a bank” did not take a substantial step towards robbing the bank).

138. See, e.g., N.Y. PENAL LAW § 140.35 (McKinney 2010) (criminalizing the possession of burglars’ tools); UTAH CODE ANN. § 76-4-401 (LexisNexis 2012 & Supp. 2013) (defining the crime of enticing a minor).

finders no longer have to deduce what the defendant would have done in order to punish him.¹³⁹ Rather, they can assign liability based on direct evidence.¹⁴⁰ In fact, numerous scholars have taken a view that the practical effect of offences such as possession of burglars' tools is to "shift the burden" to defendants regarding the nature of their activity.¹⁴¹ In other words, they might serve as an antidote to the anti-inference bias in criminal attempt cases.

To be sure, while our analysis highlights a possible explanation for the prevalence of preparatory offences, it does not necessarily justify them. Preparatory offences might effectively overcome the anti-inference bias, yet they might unjustly intrude on individual liberty,¹⁴² and, given their wide scope, they might deter socially beneficial behavior.¹⁴³ The current discussion is therefore closely tied to the discussion on debiasing. Inasmuch as legal systems are better able to deal with the anti-inference bias, the need for preparatory offences is expected to diminish. Other things being equal, a shift from the current regime—of broad prohibitions coupled with biased fact finders that are reluctant to impose liability—to a regime of narrow prohibitions that are enforced by unbiased fact finders, seems desirable. It will enable fact finders to avoid punishing blameless behavior while assuring that those who deserve to be punished will not escape liability.

C. Limitations and Future Research

In this final subpart, we acknowledge potential criticisms to this project, and outline potential avenues of future research that could help address these limitations and further our understanding of the anti-inference bias.

As is always the case with experimental work, one should be cautious regarding the external validity of our findings. A courtroom is not a decision-making lab, and

139. To be sure, such laws generally require a state of mind relating the completion of the underlying offence. *See* N.Y. PENAL LAW § 140.35. Nonetheless, it is widely recognized that such prohibitions widen the scope of liability for inchoate crimes. *See* *State v. Ansari*, 100 P.3d 231, 237–38 (Utah Ct. App. 2004) (analyzing the Utah Internet enticement law as aimed towards broadening the scope of liability); Kimberly Kessler Ferzan, *Inchoate Crimes at the Prevention/Punishment Divide*, 48 SAN DIEGO L. REV. 1273, 1283 (2011) (viewing preparatory offenses as a way in which legislature broaden the domain of liability).

140. As in the context of legal presumptions, *see supra* note 124, we do not argue that the anti-inference bias is the sole explanation for the existence of such crimes. For additional arguments, *see*, for example, Zachary Price, *The Rule of Lenity as a Rule of Structure*, 72 FORDHAM L. REV. 885, 913 (2004) (tying between the phenomenon and prosecutorial discretion); Jeannie Suk, *Criminal Law Comes Home*, 116 YALE L.J. 2, 17–19 (2006) (tying between the phenomenon and the difficulties to present evidence in court).

141. *See* Ira P. Robbins, *Double Inchoate Crimes*, 26 HARV. J. LEGIS. 1, 97 (1989); *see also* ANDREW ASHWORTH, PRINCIPLES OF CRIMINAL LAW 467 (2d ed. 1995).

142. MICHAEL MOORE, PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW 784 (1997) (noting that such crimes "give[] liberty a strong kick in the teeth right at the start"); Carissa Byrne Hessick, *Disentangling Child Pornography from Child Sex Abuse*, 88 WASH. U. L. REV. 853, 884 (2011) (arguing that such crimes "run directly counter to notions of due process and fairness in the criminal justice system").

143. *See, e.g.*, Daniel Ohana, *Responding to Acts Preparatory to the Commission of a Crime: Criminalization or Prevention?*, 25 CRIM. JUST. ETHICS 23, 28 (2006) (noting the chilling effect of preparatory offences).

different forces might interact with the anti-inference heuristic during judicial proceedings. In this regard, future research should study decisions made by both judges and jurors, and explore the extent to which the bias drives actual rulings. While exploring this question in a quantitative manner might prove a difficult task, additional methodologies such as systematic qualitative studies of jurors might prove useful.

Future research should also pay special attention to the role of two key figures that can play a critical role with respect to the anti-inference bias. Arguably, given their role as repeat players, lawyers are well aware of the precise scope of the bias. Thus, defense lawyers might attempt to strengthen the bias in order to make a favorable outcome more probable.¹⁴⁴ Framing evidence as *merely* circumstantial seems to be a routine tactic of the defense.¹⁴⁵ These lawyers might exacerbate the anti-inference bias by pointing to the existence of alternative explanations for the circumstantial evidence. As demonstrated by Niedermeier and his colleagues, a defense lawyer's argument that highlights alternative causal possibilities can be an effective way to bring about a significant decline in the inclination of fact finders to assign liability.¹⁴⁶ Furthermore, by making those alternative scenarios as vivid as possible—for instance, by using computer simulations—lawyers can render this strategy even more effective.¹⁴⁷ Mapping out the precise way defense lawyers use the bias could shed more light on the intricate details of the bias. Similarly, studying the way in which prosecutors and plaintiff lawyers attempt to overcome those efforts might point towards effective debiasing techniques.

A second group of players that merits specific research are judges. While there is no reason to assume that the psychology of legally trained subjects, including judges, differs fundamentally from that of other people—and some studies have indeed demonstrated that it does not¹⁴⁸—some studies do show that professionals can learn to employ debiasing mechanisms.¹⁴⁹ More fundamentally, one must be exceedingly cautious in deriving normative conclusions from laboratory results, as it is unclear to what extent the anti-inference bias persists in the face of legal

144. For a general analysis of the connection between jurors' decision-making patterns and the type of arguments advocates make, see Neal R. Feigenson, *The Rhetoric of Torts: How Advocates Help Jurors Think About Causation, Reasonableness, and Responsibility*, 47 HASTINGS L.J. 61 (1995).

145. Cf. Bergman, *supra* note 9, at 985 (describing lawyers in movies).

146. Niedermeier et al., *supra* note 38, at 539–41 (reporting the results of Experiment 3).

147. See, e.g., Meghan A. Dunn, Peter Salovey & Neal Feigenson, *The Jury Persuaded (and Not): Computer Animation in the Courtroom*, 28 LAW & POL'Y 228 (2006) (describing an experimental study of the effects of computer-animated displays on mock jurors); Saul M. Kassin & Meghan A. Dunn, *Computer-Animated Displays and the Jury: Facilitative and Prejudicial Effects*, 21 LAW & HUM. BEHAV. 269 (1997) (same).

148. See, e.g., Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 784–816 (2001) (reporting the results of experiments showing that federal magistrate judges are somewhat less susceptible to some cognitive illusions than lay people, yet their judgment is affected by cognitive illusions producing systematic errors).

149. Cf. Chris Guthrie & Jeffrey J. Rachlinski, *Insurers, Illusions of Judgment & Litigation*, 59 VAND. L. REV. 2017, 2222 (2006) (suggesting that insurers might have developed cognitive skills that enable them to avoid many common errors during the litigation process).

arguments aimed to counteract it. Further research involving debiasing techniques and research with professional judges would enhance our understanding of judicial behavior and the need and prospects for regulation in this context.

Finally, additional research should focus on the precise psychological mechanisms underlying the anti-inference bias. For example, even if neither the objective probability that a certain event occurred, nor the subjective assessment of that probability, differ between direct-evidence and the inference conditions, perhaps subjects' *confidence* in their subjective assessment differs between the two conditions. To be sure, a positive correlation between subjects' confidence in their probability assessment and their reluctance to impose liability based on circumstantial evidence would not ipso facto justify the anti-inference bias. It may, however, open up interesting possibilities regarding the relationship between this bias and other phenomena such as ambiguity aversion.¹⁵⁰ It may also shed light on the normative argument made in the context of naked statistical evidence, namely that fact finders should take into account not only the probability that a certain event happened, but also the weight or resiliency of the evidence supporting this probability.¹⁵¹ Perhaps one's confidence in one's probability assessment is correlated with the "weight" of the available evidence.

IV. CONCLUSION

The experiments described in this study demonstrate that the disinclination of fact finders to convict defendants or impose civil liability based solely on circumstantial evidence is more powerful than previously realized. It characterizes not only statistical, probabilistic, inconclusive, forensic, and non-case-specific evidence, but also non-statistical, non-probabilistic, conclusive, eyewitness, and case-specific evidence. None of the theories proposed as explanations for this phenomenon provide a satisfactory account, indicating the presence of a robust, deep-rooted anti-inference bias. This tendency can explain some features of extant procedural and substantive rules, and it calls for careful consideration of possible legal reforms.

Sherlock Holmes famously says that "when you have eliminated all which is impossible, then whatever remains, however improbable, must be the truth."¹⁵² Our findings indicate, however, that most people—including judicial fact finders—adopt the maxim used as the title for one of Carter Dixon's Sir Henry Merrivale mysteries, "Seeing is Believing."¹⁵³

150. While *risky* events have a certain probability for a given outcome, in *ambiguous* events even the probability of a given outcome is unknown. Ambiguity aversion is usually demonstrated through the Ellsberg paradox: people prefer to bet on an urn with fifty red and fifty blue balls, rather than on an urn with 100 balls where the number of blue and red balls is unknown. See Daniel Ellsberg, *Risk, Ambiguity, and the Savage Axioms*, 75 Q.J. ECON. 643, 650–53 (1961). Possibly, a lower confidence in one's probability assessment makes one more reluctant to act on the basis of her assessment, as in ambiguity aversion.

151. See *supra* notes 30–31 and accompanying text.

152. Arthur Conan Doyle, *The Adventure of the Blanched Soldier*, in THE CASE-BOOK OF SHERLOCK HOLMES 33, 54 (House of Stratus 2001) (1927).

153. CARTER DICKSON, SEEING IS BELIEVING (1941).