
**A Third View of the Black Box: Cognitive Coherence
in Legal Decision Making
(71 UNIV. OF CHICAGO LAW REV. 511 (2004))**

Dan Simon

USC Public Policy Research Paper No. 04-10

**PUBLIC POLICY
RESEARCH PAPER SERIES**

University of Southern California Law School
Los Angeles, CA 90089-0071

*This paper can be downloaded without charge from the Social Science Research Network
electronic library at <http://ssrn.com/abstract=541263>*

A Third View of the Black Box: Cognitive Coherence in Legal Decision Making

Dan Simon†

This Article presents a novel body of research in cognitive psychology called coherence-based reasoning, which has thus far been published in journals of experimental psychology. This cognitive approach challenges the stalemated conflict between the Rationalist and Critical models of decision making that have dominated legal scholarship for over a century. The experimental findings demonstrate that many legal decisions fit into neither of these models. Based on a connectionist cognitive architecture, coherence-based reasoning shows that the decision-making process progresses bidirectionally: premises and facts both determine conclusions and are affected by them in return. A natural result of this cognitive process is a skewing of the premises and facts toward inflated support for the chosen decision. The Article applies this research to four important aspects of the trial. It argues that the current doctrine in these areas is based on misconceptions about human cognition, which lead to systematic legal errors. By identifying the cognitive phenomena that lie at the root of these failings, the research makes it possible to devise interventions and introduce procedures that reduce the risk of trial error.

INTRODUCTION

How do judges and jurors decide cases? Though obviously central to the law, the mental processes for making decisions remain an opaque feature at the heart of legal discourse. For more than a century, views of the process have clustered around two ideal types. One rests on the assumption that legal decisions are the product of rational decision-making processes. According to this “Rationalist” view, legal decisions emanate naturally from prescribed forms of logical inference, namely deductions, inductions, and analogies. Critics, on the other hand, question the veracity of this account and portray the deci-

† Associate Professor, University of Southern California Law School. This Article and the underlying experimental research were supported by the National Science Foundation, grants SES-0080424 and SES-0080375, and by a USC Law School summer research stipend. I thank my collaborators, Keith Holyoak and Steve Read, for the stimulating and gratifying path we have taken together. For thoughtful comments on this Article and on much of the underlying research, I thank Scott Altman, Jody Armour, Yochai Benkler, Gary Blasi, Mary Dudziak, Catherine Fisk, Sam Fraidin, Ron Garett, Beth Garrett, Ariela Gross, Dan Kahan, Greg Keating, Dan Klerman, Russell Korobkin, Tom Lyon, Andrei Marmor, Ed McCaffery, Tamara Piety, Elyn Saks, Mike Shapiro, Larry Simon, Dave Slawson, Nomi Stolzenberg, and Eric Talley. For superb research, I thank Aaron Brownstein, Dan Krawczyk, Lien Pham, Adam Scardina, Chad Snow, and Carey Stone. As always, the staff of the USC Law School Library has been terrifically helpful. I owe special thanks to my mentors, Dan Gilbert, Martha Minow, Lew Sargentich, and Henry Steiner. Correspondence welcome: dsimon@law.usc.edu.

sion-making process as fundamentally inconsistent with these logical forms of inference. This alternative position, associated with Oliver Wendell Holmes, Jr. and the Legal Realists, contends that “the life of the law” is based not on logic, but rather that “the felt necessities of the time,” avowed and unconscious intuitions of public policy, and even judicial prejudices have more to do with legal decisions than the formal axioms of logical inference.¹

Few participants in contemporary legal discourse abide explicitly by either the Rationalist or the Critical conception of decision making.² In practice, however, these conceptions continue to pervade the discourse. Numerous legal practices and institutions are based on Rationalist assumptions,³ and are criticized in those terms.⁴ The dichotomous nature of the debate is unsettling in that it denies the possibility that legal reasoning might comply with neither characterization. The very fact that this debate has persisted for so long gives reason to suspect that neither account is completely convincing.

This Article introduces into the debate an emerging body of research called *coherence-based reasoning*.⁵ This experimentally based

¹ See Oliver Wendell Holmes, Jr., *The Common Law* 1 (Little, Brown 1881).

² “To some extent, we are all Legal Realists now.” Joseph William Singer, *Legal Realism Now*, 76 Cal L Rev 465, 467 (1988), reviewing Laura Kalman, *Legal Realism at Yale: 1927–1960* (North Carolina 1986). See also Gary Peller, *The Metaphysics of American Law*, 73 Cal L Rev 1151, 1152 (1985) (discussing the development of American legal thought from formalism to realism).

³ Providing specific citations to the language of logic would be beside the point given its ubiquitous presence in legal opinions and scholarship. For an exemplary instructional text, see Ruggero J. Aldisert, *Logic for Lawyers: A Guide to Clear Legal Thinking* (National Institute for Trial Advocacy 3d ed 1997).

⁴ See, for example, Steven D. Smith, *Believing Like a Lawyer*, 40 BC L Rev 1041, 1120 (1999) (describing lawyers’ “puzzling . . . combination of skeptical sophistication and apparent naïveté” in claiming that we are all legal realists while continuing to argue in modes of formalism); Pierre Schlag, *Law and Phrenology*, 110 Harv L Rev 877, 902–06 (1997) (describing the persistence of the Langdellian paradigm in contemporary jurisprudence); Paul Gewirtz, *On “I Know It When I See It,”* 105 Yale L J 1023, 1042 (1996) (stating that the typical opinion is “marked by a rhetoric of certainty, inevitability, and claimed objectivity”); J.M. Balkin, *Some Realism about Pluralism: Legal Realist Approaches to the First Amendment*, 1990 Duke L J 375, 385 (noting that First Amendment jurisprudence has only recently embraced legal realism); Lawrence Lessig, *The Regulation of Social Meaning*, 62 U Chi L Rev 943, 946 n 12 (1995) (agreeing with Balkin that First Amendment jurisprudence is predominantly formalistic); Gerald B. Wetlaufer, *Rhetoric and Its Denial in Legal Discourse*, 76 Va L Rev 1545, 1562 (1990) (stating that whenever possible, judicial opinions take the form of “deductive, syllogistic proofs”); Richard A. Posner, *The Jurisprudence of Skepticism*, 86 Mich L Rev 827, 865 (1988) (stating that even in tough cases, most judicial opinions “depict the process of reasoning as a logical deduction”); Richard A. Wasserstrom, *The Judicial Decision: Toward a Theory of Legal Justification* 16–17 (Stanford 1961) (stating that judges continue to make decisions “dictated by prior rules applied in accordance with canons of formal logic”).

⁵ The bulk of this research has been performed by Keith Holyoak, Steve Read, and me. See Keith J. Holyoak and Dan Simon, *Bidirectional Reasoning in Decision Making by Constraint Satisfaction*, 128 J Exper Psych: Gen 3 (1999); Dan Simon, et al, *The Emergence of Coherence over the Course of Decision Making*, 27 J Exper Psych: Learning, Memory & Cognition 1250

cognitive psychological theory offers a third view of the black box: a new approach to decision making that indicates a way out of the stagnant dialectic. Coherence-based reasoning posits that the mind shuns cognitively complex and difficult decision tasks by reconstructing them into easy ones, yielding strong, confident conclusions. The research reveals an unconscious transformation of the way decisions are mentally represented, ultimately leading to a seemingly straightforward choice between a compelling alternative and a weak one. To date, experimentation has revealed eight cognitive features of coherence-based reasoning, which harbor interesting—and potentially troubling—implications for decision making in the legal domain and beyond.

This Article has two objectives: to present the experimental research on coherence-based reasoning to a legal readership, and to apply it to four important procedures and practices by which we conduct and review trials. I argue that there are deep incompatibilities between actual legal decision making and the primarily Rationalist assumptions on which trials are designed. These incompatibilities render some types of legal decisions susceptible to a considerable risk of error. At the same time, since the decision-making process is also incompatible with the Critical view, we need not be as frustrated with, and distrusting of, the legal system as the Critics suggest. The research indicates that, by identifying the cognitive phenomena that lie at the root of the failings, we can devise interventions and introduce procedures that reduce the risk of error and thus make the decision-making process better fit the legal ideals it is intended to serve. This constructive endeavor is the principal aspiration of this Article. Deviation from Rationalist ideals does not necessarily imply consciously strategic or deceitful conduct on the part of the decision-maker. The research shows that these deviations can be a natural product of the cognitive

(2001); Dan Simon and Keith J. Holyoak, *Structural Dynamics of Cognition: From Consistency Theories to Constraint Satisfaction*, 6 *Personality & Soc Psych Rev* 283 (2002); Stephen J. Read, Chadwick J. Snow, and Dan Simon, *Constraint Satisfaction Processes in Social Reasoning*, in Richard Alterman and David Kirsh, eds, *Proceedings of the 25th Annual Conference of the Cognitive Science Society* 964, online at <http://homepage.psy.utexas.edu/homepage/staff/Parlee/confproc/bost03/cogsci03/index.html> (visited Apr 4, 2004).

Additional research is in the process of publication. See Dan Simon, Daniel C. Krawczyk, and Keith J. Holyoak, *Construction of Preferences by Constraint Satisfaction*, *Psych Sci* (forthcoming 2004) (on file with author); Aaron L. Brownstein, Stephen J. Read, and Dan Simon, *Effects of Individual Expertise and Task Importance on Pre-decision Reevaluation of Alternatives*, *Personality & Soc Psych Bull* (forthcoming 2004) (on file with author); Dan Simon, Chadwick J. Snow, and Stephen J. Read, *The Redux of Cognitive Consistency Theories: Evidence Judgments by Constraint Satisfaction*, *J Personality & Soc Psych* (forthcoming 2004); Aaron L. Brownstein, Dan Simon, and Stephen J. Read, *The Effects of Cognitive Constraint Satisfaction on Betting Decisions* (in progress) (on file with author); Dan Simon, Chadwick J. Snow, and Stephen J. Read, *The Self-Serving Bias Up Close* (under preparation).

process even when the decision-maker is honestly motivated to make a good decision.

One way to understand the distinction between the Rationalist and Critical positions is to compare their conceptualizations of the term “reasoning.” As the term is usually used in legal discourse, reasoning stands for the articulation of the rationale for a decision. The term can also be used to denote the actual processes of inference, that is, the mental processes by which people advance from one state of belief to the next. While the Rationalist view implicitly equates the two senses of the term, the Critical view drives a wedge between them, asserting that legal decisions are driven by factors that are incompatible with or unrelated to the espoused premises, analogies, and facts.⁶

At a deeper theoretical level, the difference between the Rationalist and Critical approaches can be understood as hinging on the *directionality* of the reasoning process. By definition, logical forms of inference follow particular courses of reasoning: deductions progress from major and minor premises toward conclusions; inductions flow from empirical observations toward generalized rules; analogies emanate from established cases and extend to target cases; and factual determinations proceed from evidence, through inferences, toward conclusions. A core tenet of the Rationalist view is that decision-making processes strictly follow these paths of inference. Critics, on the other hand, insist that the inferences are influenced, swayed, or obstructed by nonprescribed factors or, worse yet, that they proceed in the reverse (and perverse) course—creeping backward, from desired conclusions to their putative sources.

Much rides on this debate. On the Rationalist view, logical inference promotes decision making that accords with prevailing legal precepts and factual truths, thereby serving the democratic ideal of fair, legitimate, and even-handed application of law. By contrast, the Critics question the legitimacy of legal decision making, viewing it as propelled by ulterior motives, or driven by hidden biases or other fundamentally flawed forms of inference.⁷

⁶ John Dewey protested that the trouble with this syllogism is that while it “sets forth the *results* of thinking, it has nothing to do with the *operation* of thinking.” John Dewey, *Logical Method and Law*, 10 Cornell L Q 17, 22 (1924). He spoke also of the “absurd because impossible proposition that every decision should flow with formal logical necessity from antecedently known premises.” *Id.* at 25.

⁷ A frequent topic of criticism is the influence of ideology on judging. For two recent exchanges between critics and judges, compare Emerson H. Tiller and Frank B. Cross, *A Modest Proposal for Improving American Justice*, 99 Colum L Rev 215 (1999) (asserting that the federal justice system could be improved by considering the “partisan component of judging” in the process of assigning judges); Frank B. Cross and Emerson H. Tiller, *Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals*, 107 Yale L J 2155, 2168–76 (1998) (presenting data indicating that partisanship affects appellate court decisions); Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 Va L Rev 1717

This debate is not unique to legal discourse. The Rationalist view resonates with theories of logic,⁸ rational choice models of decision making,⁹ and Bayes Theorem.¹⁰ The Critics' position echoes a range of empirical work that includes cognitive research on reasoning,¹¹ the confirmation bias,¹² some aspects of biases and heuristics,¹³ the "error paradigm" in social psychology,¹⁴ and to some extent, cognitive dissonance theory.¹⁵ According to these critical bodies of research, the reasons decision-makers give—to others and possibly also to themselves—fail to correctly reflect the decision-making process; the reasons are essentially justificatory, rather than expository.

Coherence-based reasoning offers an empirical alternative to the Rationalist and Critical approaches. This account does not deny that human cognition is capable of performing some reasoning tasks in a

(1997) (claiming that ideology influences judging and presenting an empirical analysis), with Patricia M. Wald, *A Response to Tiller and Cross*, 99 Colum L Rev 235 (1999) (challenging Tiller and Cross's conclusion that judges vote ideologically); Harry T. Edwards, *Collegiality and Decision Making on the D.C. Circuit*, 84 Va L Rev 1335 (1998) (denying the influence of ideology on judging and contesting the empirical findings of Revesz's and Tiller and Cross's articles).

⁸ See, for example, Irving M. Copi and Carl Cohen, *Introduction to Logic* (Prentice Hall 11th ed 2001).

⁹ See, for example, John Von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior* 1–45 (Princeton 2d ed 1947).

¹⁰ For more on Bayes Theorem, see Part II.B.

¹¹ For example, Evans and Wason have shown that people's justifications of their reasoning in performing logical tasks do not reflect insight into the underlying thought processes, but are simply plausible accounts that seem to provide a good explanation for their conclusions. See Jonathan St. B.T. Evans and P.C. Wason, *Rationalization in a Reasoning Task*, 67 Brit J Psych 479, 483–86 (1976). See also Jonathan St. B.T. Evans and David E. Over, *Rationality and Reasoning* 109–17 (Psychology 1996) (discussing "belief bias," that is, the tendency to judge the validity of reasoning according to one's a priori beliefs about the conclusion of that reasoning).

¹² Confirmation bias concerns a general tendency to inappropriately bolster a hypothesis or conclusion whose truth is in question, due to preexisting beliefs about that hypothesis or conclusion. For a comprehensive review, see Raymond S. Nickerson, *Confirmation Bias: A Ubiquitous Phenomenon in Many Guises*, 2 Rev Gen Psych 175 (1998).

¹³ One example is hindsight bias. Experimental findings demonstrate that responses after an event are influenced by knowledge of the outcome of the event (which, of course, was not known prior to the event). The ensuing judgments give one the erroneous sense that they were justified by the ex ante information. See, for example, Scott A. Hawkins and Reid Hastie, *Hindsight: Biased Judgments of Past Events after the Outcomes Are Known*, 107 Psych Bull 311 (1990). For a rich demonstration of applications of biases and heuristics, see Edward J. McCaffery and Jonathan Baron, *Thinking about Tax* (draft Mar 12, 2004) (on file with author).

¹⁴ On the cognitive origins of erroneous beliefs, see, for example, Ziva Kunda, *Social Cognition: Making Sense of People* 15–52 (MIT 1999); Thomas Gilovich, *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life* 1–6 (Free Press 1993).

¹⁵ The typical experimental paradigm of cognitive dissonance theory is "forced compliance." In these experiments, participants are induced to perform an act that runs counter to their initial attitudes or moral standards. The consistent finding is that participants change their attitudes and beliefs to comport with their behavior. Newly adopted attitudes and beliefs provide a justification for past conduct. On the current state of this research, see Eddie Harmon-Jones and Judson Mills, eds, *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology* (American Psychological 1999).

manner consistent with the Rationalist account, nor that people violate these assumptions in some circumstances. Rather, it proposes a theory of cognition that contains elements of both approaches. A central tenet of the theory is that decisions are the product of a cognitive mechanism that operates *bidirectionally*, both in the prescribed and the reverse directions of reasoning.

Coherence-based reasoning applies to mental tasks in which the person must make a discrete decision or judgment in the face of complexity. Tasks are said to be complex when their constitutive considerations are numerous, contradictory, ambiguous, and incommensurate.¹⁶ Most legal cases that are litigated and appealed are of this nature, in that the facts can be ambiguous, incomplete, and contradictory; different rules, values, and principles can be invoked to support opposite conclusions; and the case at hand can be somewhat analogous to more than one previous decision.¹⁷ On their face, such tasks might seem intractable.

In a nutshell, coherence-based reasoning suggests that decisions are made effectively and comfortably when based on *coherent mental models*. Loosely defined, mental models capture the decision-maker's perception of the task at hand—that is, the way the considerations of the decision are represented in her mind.¹⁸ A mental model of a decision task is deemed “coherent” when the decision-maker perceives the chosen alternative to be supported by strong considerations while the considerations that support the rejected alternative are weak.¹⁹ Such is the case, for example, when the prosecution's eyewitness is reliable, the forensic evidence is compelling, and the defendant has a strong motive and a weak alibi. A mental model is considered “incoherent” when the decision-maker perceives the considerations as providing equivocal support for both alternatives. As defined, coherence is an empirical phenomenon, not a jurisprudential ideal.²⁰

¹⁶ William James described decision tasks in a similar fashion: “many objects, purposes, reasons, motives, [are] related to each other, some in a harmonious and some in an antagonistic way.” William James, 2 *The Principles of Psychology* 1136 n 39 (Harvard 1981) (originally published in 1890).

¹⁷ See Frederick Schauer, *Judging in the Corner of the Law*, 61 S Cal L Rev 1717, 1723–26 (1988) (arguing that disputes that lead to litigation tend to be more novel and indeterminate in their interpretive possibilities than disputes that do not lead to litigation).

¹⁸ The term mental model is used here in the broad sense of a structured representation. See Arthur B. Markman, *Knowledge Representation* 248–76 (Lawrence Erlbaum 1999).

¹⁹ As in most models of decision making, decisions are understood to be determined by the level of support they receive from the task variables. This tradition goes back to William James. James noted that “[t]he reinforcing and inhibiting ideas meanwhile are termed the *reasons* or *motives* by which the decision is brought about.” James, 2 *Principles of Psychology* at 1136 (cited in note 16).

²⁰ Compare Ronald Dworkin, *Law's Empire* 225–26 (Belknap 1986) (advocating coherence in the relationship between a proposition entailed in a judicial decision and precepts of law that are exterior to the decision itself).

The central finding of coherence-based reasoning research is that the cognitive system *imposes* coherence on complex decision tasks. Throughout the decision-making process, the mental representation of the considerations undergoes gradual change and ultimately shifts toward a state of coherence with either one of the decision alternatives. Due to these *coherence shifts*, at the culmination of the process, the decision-maker's mental model is skewed toward conformity with the emerging decision. As the hard case morphs into an easy one, the decision follows easily and confidently. The fact that decisions are ultimately based on skewed mental models and backed by high levels of confidence facilitates the making of the decision, but at the same time it can also harbor problematic implications.

Coherence shifts are a robust finding. As described below, they have been consistently observed in experiments using a number of different stimulus materials. Shifts have been observed in evaluations of factual inferences, premises, analogies, preferences, background beliefs, public policy issues, and more. Despite the ambiguity and complexity of the cases, the perceptions of the cases always shift toward either one of the two polarized mental models.

Coherence-based reasoning presents a general model of judgment and decision making in conditions of complexity. The theory parallels, but differs from, the biases and heuristics research, which underlies the burgeoning field of behavioral economics.²¹ These two bodies of research examine different cognitive phenomena. The biases and heuristics literature deals primarily with a relatively loose assortment of specific and narrowly defined tasks of judgment and choice, and does not purport to offer a cohesive or general psychological theory.²² Coherence-based reasoning is concerned with the underlying cognitive mechanism that drives the processing of large and complex processing tasks.

Coherence-based reasoning bears some resemblance to the research on group polarization, which has been successfully applied by Cass Sunstein to decision making in various legal contexts.²³ Group

²¹ For a recent review of the research on biases and heuristics, see Thomas Gilovich, Dale Griffin, and Daniel Kahneman, eds, *Heuristics and Biases: The Psychology of Intuitive Judgment* (Cambridge 2002). For applications of this body of work to law, see Cass R. Sunstein, ed, *Behavioral Law & Economics* (Cambridge 2000); *Symposium: Empirical Legal Realism: A New Social Scientific Assessment of Law and Human Behavior*, 97 Nw U L Rev 1075 (2003); Russell B. Korobkin and Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 Cal L Rev 1051 (2000).

²² Biases and heuristics seem to be part of the brain's tendency to conserve mental effort, but that is too general a function to make for an operative psychological theory. For a proposed conceptualization of the field, see Mark Kelman, *Law and Behavioral Science: Conceptual Overviews*, 97 Nw U L Rev 1347 (2003).

²³ See Cass R. Sunstein, *Deliberative Trouble? Why Groups Go to Extremes*, 110 Yale L J 71 (2000) (discussing evidence and implications of group polarization in legal and political insti-

polarization research shows that deliberation in groups causes systematic shifts that amplify the group members' predeliberation positions;²⁴ coherence-based reasoning reveals a similar amplification at the intrapersonal level. There is good reason to believe that the two phenomena interact with and compound one another.

The bidirectional interaction between the underlying considerations and the ultimate decision bears a structural resemblance to John Rawls's concept of reflective equilibrium.²⁵ Despite the structural resemblance, however, there is an important difference between reflective equilibrium and coherence-based reasoning. The former is a normative approach and is prescribed as a process best suited to solve a category of judgments. Coherence-based reasoning, on the other hand, is an empirical construct. It is a nonconscious cognitive process that drives decisions of a certain type, and causes both sanctioned and non-sanctioned changes in the mental models representing the task.

Part I of this Article introduces the cognitive theory underlying coherence-based reasoning and reviews some of the experimental findings that support it. Together, these findings amount to an elaborate account of cognitive processing in conditions of complexity.

In Part II, I apply the research to four existing procedures and practices used to conduct and review trials. The first application follows from the finding that coherence shifts occur prior to the making of a decision. This suggests the possibility that, in both criminal and civil trials, jurors integrate the evidence into coherent mental models before having been instructed on the law. If the jurors' preconceptions of the applicable law are incorrect, this could readily lead to legally in-

tutions); Cass R. Sunstein, *Republic.com* 71–80 (Princeton 2001) (discussing group polarization with respect to the Internet, media proliferation, and group isolation).

²⁴ See Daniel J. Isenberg, *Group Polarization: A Critical Review and Meta-analysis*, 50 *J Personality & Soc Psych* 1141 (1986); David G. Myers and Helmut Lamm, *The Group Polarization Phenomenon*, 83 *Psych Bull* 602 (1976).

²⁵ In searching for a principled judgment, Rawls advocates working “from both ends,” gradually revising and harmonizing the conception of the initial condition and the judgment derived from it. Rawls explains:

By going back and forth, sometimes altering the conditions of the contractual circumstances, at others withdrawing our judgments and conforming them to principle, I assume that eventually we shall find a description of the initial situation that both expresses reasonable conditions and yields principles that match our considered judgments duly pruned and adjusted. This state of affairs I refer to as reflective equilibrium.

John Rawls, *A Theory of Justice* 18 (Belknap rev ed 1999). The process is as much about justifying one's particular judgment as it is about reconceptualizing the principles of the original position. At equilibrium, the justification of a conception of justice is a function of the “mutual support of many considerations, of everything fitting together into one coherent view.” *Id.* at 19. The idea that reflective equilibrium can drive an epistemological or ethical model is inimical to foundationalist philosophical theories. The latter tend to depend critically on the unidirectionality of the reasoning process, from foundational premises to inferred beliefs and judgments. See Paul Thagard, *Coherence in Thought and Action* 4–5 (MIT 2000).

correct decisions. I propose introducing substantive instructions at a preliminary stage of the trial, and discuss various implications of the proposal.

Second, coherence-based reasoning speaks to an ongoing debate as to whether fact-finders evaluate evidence in a holistic or an atomistic manner. This debate has particular significance with respect to the admissibility of potentially prejudicial evidence. While the empirical findings support the holistic account, I argue that its corollary prescription is wrong. I also caution against a Supreme Court decision that tends toward a relaxation of the restraints on the admissibility of evidence.²⁶

The third application concerns verdict determinations in criminal trials. The primary finding of the research is that coherence shifts polarize perceptions of the evidence. Jurors with a slight initial inclination to acquit or convict are likely to amplify their perception of the case, so that evidence that is weakly probative of guilt can be transformed to create mental models that strongly support either innocence or guilt. Normatively, this matters for jurors who vote to convict, because the evidence is bolstered from overall ambiguity to a belief beyond a reasonable doubt. Thus, coherence shifts can undermine the role of the heightened standard of proof. Initial research shows that this problem can be moderated by means of a fairly straightforward instruction aimed at neutralizing the bias. I propose that, contingent on further research, juries be given such an instruction.

The final application pertains to the doctrine of harmless error in criminal trials, which has been torn between two modes of analysis: one that focuses on assessing the impact of the error on the trial outcome, and the other that focuses on assessing guilt as it appears from the remaining, presumably untainted, evidence. The Supreme Court tends to apply the latter, guilt-focused approach. The context of harmless-error review is unique in that it gives judges considerable latitude in structuring their decision. In cases that concern the erroneous admission of impermissible evidence, the appellate judge's perception of the remainder of the evidence is likely to have been affected by the impermissible evidence due to coherence effects. Since that evidence might well be tainted, judges ought to shy away from guilt-focused analysis. For cases that involve errors other than impermissible evidence, coherence-based reasoning does not by itself warrant a strong prescription, but it does offer some reasons for preferring error-based analysis.

In the Conclusion, I claim that coherence-based reasoning applies to a considerably broader range of legal issues than the four examples

²⁶ See *Old Chief v United States*, 519 US 172 (1997).

discussed in Part II. To date, the experimental research is moving faster than the applications to law, and with more research underway, the theory harbors immense potential for providing unique insight to many more legal applications.

I. COHERENCE-BASED REASONING: THEORY AND EXPERIMENTAL SUPPORT

This Part of the Article presents the psychological underpinnings of coherence-based reasoning. After a short introduction to the key theoretical aspects, I provide an account of the experimentation in some detail. The research necessary for the legal applications in Part II is summarized and discussed at the end of this Part.

A. The Cognitive Theory

Connectionist Representations. All mental processing starts with a mental representation of the task.²⁷ Decisions are made within certain problem spaces that contain and delimit the representation of the task variables, that is, all the factors that are to be included in the decision at hand. The representations of the variables are called *mental models* of the decision task.

Coherence-based reasoning presumes a connectionist architecture of mental representations, which can be likened to an intricate electrical network.²⁸ Each variable is represented in the mind as a unit,

²⁷ In order to perform any mental operation, the variables involved in the task must first be represented somehow in the mind. Put succinctly, “[a]ll that you know about the world and your only basis for acting on the world is found in your mental representations.” Ronald T. Kellogg, *Cognitive Psychology* 8 (Sage 2d ed 2003). See also Lawrence W. Barsalou, *Cognitive Psychology: An Overview for Cognitive Scientists* 55 (Lawrence Erlbaum 1992).

Representation should not be understood as entailing conscious recognition. In Dennett’s terms, it need not be in one’s “speech centre.” See D.C. Dennett, *Content and Consciousness* 118–19 (Humanities 1969).

²⁸ The exact nature of mental representations is a topic of intense debate between advocates of two general approaches: connectionist models and symbolic systems. For reviews of the debate, see Keith J. Holyoak and John E. Hummel, *The Proper Treatment of Symbols in a Connectionist Architecture*, in Eric Dietrich and Arthur B. Markman, eds, *Cognitive Dynamics: Conceptual and Representational Change in Humans and Machines* 229, 229–31 (Lawrence Erlbaum 2000); Gary F. Marcus, *The Algebraic Mind: Integrating Connectionism and Cognitive Science* 1–4 (MIT 2001).

The research reported in this Article is consistent with a hybrid conception of representation, called symbolic connectionism. See Keith J. Holyoak, *Symbolic Connectionism: Toward Third-Generation Theories of Expertise*, in K. Anders Ericsson and Jacqui Smith, eds, *Toward a General Theory of Expertise: Prospects and Limits* 301 (Cambridge 1991); Holyoak and Hummel, *Proper Treatment of Symbols* at 232–59 (articulating the theoretical position for a network-like structure of symbolic knowledge).

Symbolic connectionist networks do not define variables at the neuron level. As Holyoak and Spellman explain, systematic reasoning requires symbolic representations without which links between elements would be incapable of defining meaningful relationships. See Keith J. Holyoak and Barbara A. Spellman, *Thinking*, 44 *Ann Rev Psych* 265, 270–72 (1993). Representa-

which is then connected through the network to all the other units.²⁹ The initial representation of a mental task is determined foremost by the rich and detailed store of background knowledge about one's physical, social, and conceptual worlds.³⁰ The mental model of a complex decision task contains a myriad of variables that point in more than one direction and thus do not all fit into a coherent mental model. One subset of variables $[a_1, a_2, \dots, a_n]$ supports conclusion A, and the other subset $[b_1, b_2, \dots, b_n]$ supports the opposite conclusion B.

Constraint Satisfaction Mechanisms. In all but easy cases, neither subset of variables dominates the other. Since each variable has some bearing on the task, it can be said to impose a *constraint* on the network.³¹ Connectionist systems process mental tasks through a nonconscious process in which the mentally represented variables interact with one another like an electrical network. Activation spreads throughout the mental model.³² Each and every constraint influences, and is influenced by, the entire network, so that every processing cycle results in a slightly modified mental model. Over time, unsupported variables or those suppressed by other variables degrade and even die out, while those that are mutually supported gain strength.

tions are thus constructed by high-level conceptual building blocks that share some of the symbolic properties of more traditional cognitive architectures. See Stephen J. Read, Eric J. Vanman, and Lynn C. Miller, *Connectionism, Parallel Constraint Satisfaction Processes, and Gestalt Principles: (Re)Introducing Cognitive Dynamics to Social Psychology*, 1 *Personality & Soc Psych Rev* 26, 32 (1997).

²⁹ The principal characteristic of connectionist representations is the level of activation of the elements. The level of activation stands for the degree of the respective variable's strength, or acceptability. Elements are connected to other elements by links, which can be either positive (supportive) or negative (inhibitory). Coherent relationships are denoted by positive links; inconsistent relationships are denoted by negative links. Links also vary in degree: where the association between the respective variables is strong, the link is said to carry much weight.

The seminal work on constraint satisfaction theories is David Rumelhart and James McClelland, eds, 1 *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (MIT 1986). For introductions to connectionism, see Thagard, *Thought and Action* at 1–13 (cited in note 25) (explaining the importance of coherence as a concept in philosophy and psychology); William Bechtel and Adele Abrahamsen, *Connectionism and the Mind: An Introduction to Parallel Processing in Networks* (Basil Blackwell 1991). For an excellent review of the field, see Read, Vanman, and Miller, 1 *Personality & Soc Psych Rev* 26, 27–32 (cited in note 28).

³⁰ For a review of representations of knowledge, see Markman, *Knowledge Representation* at 1–26 (cited in note 18). See also Stephen J. Read, *Constructing Causal Scenarios: A Knowledge Structure Approach to Causal Reasoning*, 52 *J Personality & Soc Psych* 288 (1987).

³¹ Each constraint is a function of the respective element's level of activation and the weight of its links. Each variable is said to constitute a soft constraint. See Bechtel and Abrahamsen, *Connectionism and the Mind* at 58–60 (cited in note 29).

³² Positively linked elements excite one another, while negatively linked ones inhibit each other. The induced activation of any pair of linked elements depends on their relative initial levels of activation and on the weight and sign of the link that connects them. Highly activated elements strongly affect others and are resistant to external influence; weakly activated elements hardly affect others and are sensitive to external activations.

The central feature of constraint satisfaction mechanisms is that the mental model will reconfigure itself until the constraints settle at a point of maximal coherence. Coherence, the state at which similarly linked variables are similarly activated, occurs when the variables that support the emerging decision are strongly endorsed and those supporting the alternative are dismissed, rejected, or ignored. In complex decisions, the initial mental representation of the task is naturally incoherent. Constraint satisfaction processes force the task variables to change toward a better fit with the gradually emerging state of coherence.³³ This reversed induction gives coherence-based reasoning its bidirectional character: while the strength of supporting variables determines the conclusion, the variables themselves are transformed by the cognitive process so as to provide considerably stronger support for the conclusion.³⁴

In sum, the ultimate state of coherence is essentially a byproduct of the cognitive system's drifting toward either one of two skewed mental models. Within each of these models, the initially complex and incoherent mental model has been spread into two subsets, one of which dominates the other, thereby enabling a relatively easy and con-

³³ The phenomena of reversed induction and structural reconfiguration are the hallmarks of Gestalt psychology. Gestalt psychologist Max Wertheimer explains "one has a feeling how successive parts should follow one another; one knows what a 'good continuation' is; how 'inner coherence' is to be achieved, etc.; one recognizes a resultant 'good Gestalt' simply by its own 'inner necessity'." Max Wertheimer, *Laws in Organization of Perceptual Forms*, in Willis D. Ellis, ed., *A Source Book of Gestalt Psychology* 71, 83 (Routledge 1955).

Constraint satisfaction theory has been dubbed a "new kind of Gestalt theory." Stephen E. Palmer, *Gestalt Psychology Redux*, in Peter Baumgartner and Sabine Payr, eds., *Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists* 160 (Princeton 1995). See also Keith J. Holyoak, *Problem Solving*, in Edward E. Smith and Daniel N. Osherson, eds., *An Invitation to Cognitive Science: Thinking* 267, 289–90 (MIT 2d ed 1995) (discussing similarities between the concepts of constraint satisfaction and Gestalt theory). The Gestaltian roots of connectionism have also been discussed in Simon and Holyoak, 6 *Personality & Soc Psych Rev* at 283–85 (cited in note 5).

³⁴ Another way to appreciate how coherence-based reasoning differs from conventional theories of cognition is to consider how it breaks with the prevalent understanding of the relationship between the two building blocks of cognitive psychology—representation and computation. At the core of cognitive psychology is the general idea that mental activity is the product of applying a procedure, computation, or manipulation to a state of affairs represented in the mind. In general terms, all cognitive processing entails the dual function of representation and computation. The latter component is best understood as a form of making an inference or judgment, whereas the former concerns mentally representing the variables. See note 28.

The dominant theories of cognitive psychology presume separability of the two components, as do most theories of decision making and legal reasoning. These theories assume invariance and exogeneity of mental representations. That means (1) that the representations of the task variables do not change during the performance of the task, at least not in any systematic manner; and (2) that the representations of the task variables are determined by factors such as background knowledge, personal preferences, cultural conventions, and the like, which are external and antecedent to the decision-making task itself. Coherence-based reasoning demonstrates that mental representations are not entirely invariant or exogenous to the process of making decisions, and suggests that representational shifts are central to the decision-making process itself.

fidient choice. This skewed representation reflects an artificial polarization between the inflated representation of the variables that support the chosen conclusion and the deflated ones that support the rejected conclusion; it differs considerably from the way the task variables were perceived before the decision-making process got underway, and it differs also from the way they will be perceived some time after the completion of the task.

While coherence-based reasoning is a nascent field of research, constraint satisfaction mechanisms have been the subject of quite active research over the last two decades.³⁵ This research, however, has been based primarily on computer modeling and has focused on the role of coherence as a criterion for judgments and decision making. My collaborators and I designed coherence-based reasoning research as an experimental paradigm, with the additional ambition of demonstrating the effects of the decision-making process on the mental models of the task variables.

B. The Experimental Research

The following Part outlines the empirical support for the eight primary features of coherence-based reasoning. The research focuses on two series of experiments: one based on the case of *Quest v Smith*, performed in collaboration with Keith Holyoak of the Department of Psychology at UCLA, and one based on the case of *Jason Wells*, per-

³⁵ Paul Thagard applied constraint satisfaction theory to the construction of scientific and factual explanations. See Paul Thagard, *Explanatory Coherence*, 12 *Behav & Brain Sci* 435 (1989). Keith Holyoak and Paul Thagard modeled analogical reasoning in a computer simulation. See Keith J. Holyoak and Paul Thagard, *Analogical Mapping by Constraint Satisfaction*, 13 *Cognitive Sci* 295 (1989). Lynn Miller and Steve Read have used computerized models of constraint satisfaction to demonstrate the role of coherence in social judgment. See Lynn Carol Miller and Stephen J. Read, *On the Coherence of Mental Models of Persons and Relationships: A Knowledge Structure Approach*, in Garth J.O. Fletcher and Frank D. Fincham, eds, *Cognition in Close Relationships* 69 (Lawrence Erlbaum 1991); Stephen J. Read and Amy Marcus-Newhall, *Explanatory Coherence in Social Explanations: A Parallel Distributed Processing Account*, 65 *J Personality & Soc Psych* 429 (1993) (describing experimental results consistent with Thagard's theory of explanatory coherence).

Constraint satisfaction has also been applied to model the core phenomena of cognitive dissonance theory, see Thomas R. Schultz and Mark R. Lepper, *Cognitive Dissonance Reduction as Constraint Satisfaction*, 103 *Psych Rev* 219 (1996) (contrasting cognitive dissonance theory with a constraint satisfaction computer model in reference to human data); decisional dilemmas, see Paul Thagard and Elijah Millgram, *Inference to the Best Plan: A Coherence Theory of Decision*, in Ashwin Ram and David B. Leake, eds, *Goal-Driven Learning* 439, 449–53 (MIT 1995) (modeling decisional dilemmas by constraint satisfaction); and stereotyping, see Ziva Kunda and Paul Thagard, *Forming Impressions from Stereotypes, Traits, and Behaviors: A Parallel-Constraint-Satisfaction Theory*, 103 *Psych Rev* 284, 300–04 (1996) (proposing that the interaction between stereotypical and individuating information can be modeled as a constraint satisfaction process); Paul Thagard and Ziva Kunda, *Making Sense of People: Coherence Mechanisms*, in Stephen J. Read and Lynn C. Miller, eds, *Connectionist Models of Social Reasoning and Social Behavior* 3 (Lawrence Erlbaum 1998) (same).

formed in collaboration with Stephen Read of the Department of Psychology at USC.³⁶

1. The basic finding: coherence shifts.

The first finding concerns the fundamental property of coherence-based reasoning: that in the course of making a complex decision, the mental model of the task shifts toward coherence with the emerging decision.

*The Case of Quest v Smith.*³⁷ Participants were first presented with a pretest that contained a number of apparently unrelated vignettes. Each vignette was followed by a statement or two that could be inferred from the vignette. Participants were asked to rate their agreement with each inference. Participants were told that they were not expected to have any expert knowledge, and were encouraged to use common sense and general knowledge in making their ratings. Ratings were made on an eleven-point scale, ranging from -5 (“strongly disagree”) to +5 (“strongly agree”), with a rating of 0 indicating neutrality. In total, participants rated twelve inferences.³⁸ Some vignettes involved factual judgments,³⁹ and some involved more abstract issues such as analogies and issues of social policy.⁴⁰ Unbe-

³⁶ The bulk of this research was funded by the National Science Foundation.

³⁷ This experiment was reported as Experiment 1 in Holyoak and Simon, 128 *J Exper Psych: Gen 3* (cited in note 5). Forty-eight undergraduate students from UCLA participated in this study to satisfy a course requirement.

³⁸ The twelve inferences were based on ten vignettes. For technical reasons, two vignettes were followed by a pair of statements rather than a single statement.

³⁹ For example, one vignette read as follows:

A company that develops computer software was a great success when it started up, but later it began to suffer from stiff competition. It did not improve its original products. Internal disputes caused the management to perform poorly. As its sales have fallen, the company has been losing money and borrowing more to fund its operations. Its stock price has been slipping steadily for months. An investor had at first made a profit from his investment in the company, but more recently has watched his investment shrivel. He discovered that the company had not been upgrading its products. The investor was becoming increasingly frustrated with the company's performance; he was especially angered by its top executive's brazen denial of any troubles and hyperbolic promises for a great future. The investor spreads a message that the company is in a mess and its situation is hopeless.

Participants were then asked to rate their agreement with the following two statements: “The investor's action was motivated primarily by vindictiveness; he was a sore loser who was happy to make money from the company's success but unwilling to sit through tough times”; and “The investor's main intention was to prevent other innocent potential investors from being misled into a bad investment.”

⁴⁰ For example, participants were presented the following text: “A newspaper is defined as: ‘any publication intended for the distribution and dissemination of news, facts, or opinions to broad audiences.’ An electronic ‘bulletin board’ is a forum on computer networks (such as the Internet) where subscribers post messages that are open to other subscribers.” They were then asked to rate their agreement with the following statement: “As a matter of legal policy, messages posted on electronic bulletin boards should be treated like items published in newspapers.”

knownst to the participants, these vignettes constituted the parts of a legal case which they would later be asked to decide.

In the second phase of the experiment, participants were asked to play the role of a young judge, assigned to decide a civil case in which Quest, a software company, filed a libel lawsuit against one of its shareholders, Jack Smith. The materials contained the summary of the evidence, which was loosely based on an actual case. The facts were not in dispute. Quest's financial situation was deteriorating and its managers were struggling to cope with the problems facing the company. Smith, a dissatisfied shareholder, posted a negative message on an electronic bulletin board directed at investors in which he sharply criticized the company's management and predicted a precipitous downfall. Shortly thereafter, Quest's stock price plummeted and the company went bankrupt. It was later revealed that Quest had been secretly developing a new product that might have saved the company, though this information was not known to Smith.

Each of the parties made six arguments, forming opposing pairs on the six points of dispute.⁴¹ Crucially, the litigants' arguments were virtually identical to the inferences that followed the vignettes in the pretest. Overall, the facts of the case were ambiguous, and the parties' arguments were plausible and balanced so as to create a complex decision.

Participants were asked to render a verdict in favor of either party. Again, they were reminded that they were not expected to have any expert knowledge, but were asked to use common sense and to base their decisions "as soundly and fairly as possible on the available facts and arguments from both sides." The verdict was followed by a request to "[r]ate your confidence that you have made the best possi-

Similarly, participants were presented the following text: "A telephone system is defined as 'a network of interconnected lines used to transmit and receive voice or data from one extension to one or more other extensions.' An electronic 'bulletin board' is a forum on computer networks (such as the Internet) where subscribers post messages that are open to other subscribers." They were then asked to rate their agreement with the following statement: "As a matter of legal policy, messages posted on electronic bulletin boards should be treated like messages sent over a telephone network."

⁴¹ The first three points of dispute involved matters of fact: Quest argued that Smith's negative message was unfounded, whereas Smith claimed it was true; Quest asserted that the message caused the company's downfall, whereas Smith claimed that mismanagement was the cause; and Quest claimed that Smith's action was motivated by vindictiveness, whereas Smith claimed he only aimed to protect other innocent investors. The remaining three points of dispute involved matters of law and social policy: Quest argued that as a matter of public policy, it is in society's interest to regulate speech over the Internet, whereas Smith argued that society would benefit more from an unregulated Internet; Quest likened the Internet to a newspaper, which according to precedent, was subject to libel law, whereas Smith drew an analogy to a telephone system, which is immune from libel law; and Quest claimed that, in posting his message, Smith had violated a company bylaw requiring prior notification of management, whereas Smith maintained that he had complied with the bylaw.

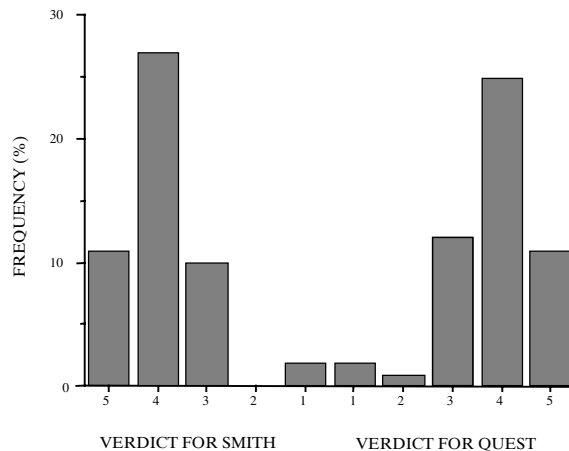
ble decision.” The confidence rating was made on a five-point scale, ranging from 1 (low), through 3 (medium), to 5 (high). Participants were then asked to rate their agreement with the twelve arguments made by the parties. These questions were essentially identical to the questions asked on the pretest.⁴² The key measure in this experiment and those that follow is the comparison between the ratings given on the vignettes with those given at the subsequent stage of making the decision. Coherence-based reasoning would predict that emerging decisions will be accompanied by a general shift in the ratings, from the initial, spontaneous state in which the variables are neither coherent with the eventual decision nor with the other variables, to a state in which they cohere with the decision and with other variables that support the same decision.

All the results were consistent with these predictions. First, we found that the participants made confident decisions despite the complexity of the case. Participants reported high levels of confidence in their decisions: 75 percent of participants indicated that they had maximal or next-to-maximal confidence in their verdicts; conversely, only 5 percent indicated low or next-to-low confidence.⁴³

⁴² For example, with regard to the second issue argued by the parties, participants were asked to state their agreement with the following statements: “Smith’s message caused Quest’s collapse”; and “Quest collapsed because of poor performance.”

⁴³ Participants were about evenly divided in their verdicts, with twenty-six deciding in favor of the plaintiff, Quest, and twenty-two deciding in favor of the defendant, Smith. Figure 1 depicts the distribution of confidence levels, plotted for all participants on a single axis, with the high values at the extreme ends of the scale. The heights of the bars represent the percentage of participants who reported the respective level of confidence.

FIGURE 1



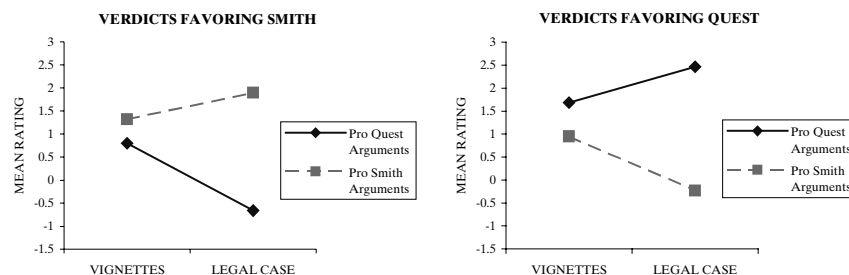
Next, we examined the crucial comparison between the ratings given at the two phases of the experiment. We found that the vignette ratings were mostly moderate, whereas by the conclusion of the legal decision, the ratings of the arguments had shifted to polarized states of coherence: participants who voted for Smith reported strong agreement with the arguments that supported the Smith verdict and disagreement with the arguments that supported Quest's position, while opposite ratings were reported by participants who decided for Quest.⁴⁴

Coherence shifts can also be observed for individual points of dispute. For example, participants were told that legally, a claim of libel can be based on an expression published in a newspaper, but not an expression communicated over the telephone. The materials contained putative legal definitions of both a newspaper and a telephone system, each of which was sufficiently ambiguous to be seen as applicable to the Internet.⁴⁵ Based on these definitions, the lawyer for the plaintiff Quest argued that the Internet is like a newspaper and thus exposed to libel law. The lawyer for defendant Smith argued that the Internet is closer to the legal definition of a telephone system and thus immune from libel law. In the post-test questionnaire, participants were asked to rate their agreement with each of these analogies.⁴⁶

⁴⁴ The left panel of Figure 2 depicts the coherence shifts for participants who voted for Smith. The vertical axis represents the mean rating with the legal arguments, plotted separately for arguments that supported Quest's position and for arguments that supported Smith's position. The combined mean of the arguments that favor the chosen decision increased from +1.32 to +1.9, whereas the combined mean of the arguments supporting the rejected decision decreased from +0.8 to -0.66.

The right panel depicts the coherence shifts for participants who voted for Quest. The combined mean of the arguments that favor the chosen decision increased from +1.68 to +2.46, whereas the combined mean of the arguments supporting the rejected decision decreased from +0.95 to -0.23.

FIGURE 2



Unless stated otherwise, all results presented in this Article were significant at the level of $p < 0.05$, and most were significant at the level of 0.001.

⁴⁵ For the definitions of a newspaper and of a telephone system, see note 40.

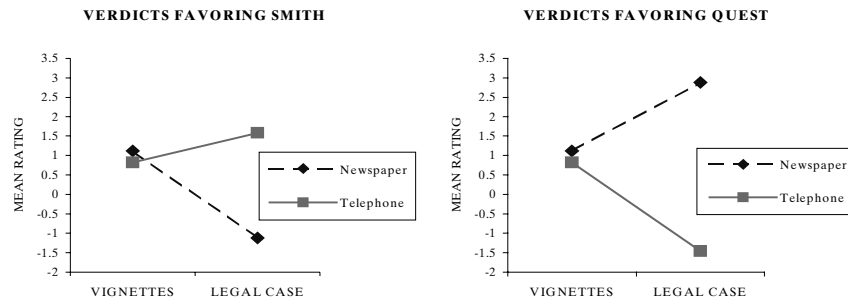
⁴⁶ Figure 3 shows that at the vignette phase, participants gave similarly positive ratings for

Coherence shifts were evident not only from the combined mean shifts of the ratings, but also from an inter-correctional analysis.⁴⁷ It is noteworthy that the coherence shifts occurred within a single experimental session of only one hour.⁴⁸

*The case of Jason Wells.*⁴⁹ The second set of materials was designed to test simulated jury decision making in a fact-laden criminal

both analogies (positive ratings were expected since both analogies were plausible). However, when the same analogies were evaluated in the context of the case, the ratings shifted in a predictable pattern: the initially moderate ratings clustered around the range from +0.82 to +1.12, and then increased to +1.58 and +2.88, and decreased to -1.12 and -1.46, all in the predicted directions.

FIGURE 3



⁴⁷ Although these analyses reveal a clear shift in the aggregate ratings of the six points of dispute, they do not suffice to establish that individual participants reached a broadly coherent position across all the disputed issues. It remains theoretically possible that these mean effects were caused by shifts in just one or two arguments rather than by imposition of coherence on the entire sets of arguments.

To rule out this possibility, we performed a correlational analysis of all six arguments and the eventual verdicts. We predicted that at the vignette phase, participants' ratings of the six positions would not constrain one another, and hence would tend to be uncorrelated, and that at the decision phase coherence effects would yield positive correlations among the disputed points, and between each point and the verdict. This, in fact, is the pattern we observed. In the first measurement, we found that only two of the twenty-one correlations were significantly positive, and several were negative. This further demonstrates that the materials in fact created a genuine sense of complexity and ambiguity. In contrast, in the second measurement, all but one of the twenty-one correlations were significantly positive, including all six correlations between disputed points and the verdict; the nonsignificant correlation was also positive.

⁴⁸ Initially, we suspected that coherence shifts might not be obtained within a single experimental session, because participants might be sufficiently aware of their initial responses and would maintain consistency in their responses. We first ran the *Quest* experiment with a week delay between the two phases. See Experiment 2, in Holyoak and Simon, 128 *J Exper Psych: Gen* at 10–11 (cited in note 5). In subsequent experiments, we found that our concerns were groundless: similar shifts occurred within a single experimental setting of one hour.

⁴⁹ This research on juror decision making is reported in Simon, Snow, and Read, *Evidence Judgments* (cited in note 5). The specific study described here is the basic study with the *Jason Wells* materials. It was ultimately left out of that article, though it overlaps with the studies reported there.

Ninety-one participants took part in this basic study. A total of 1,077 people took part in the four studies. The participants ranged in age between eighteen and eighty with a mean age of thirty-four. These experiments were completed over the Internet. Participants were solicited through links posted on a number of psychology-related websites and advertisements placed on

case. We conducted these experiments, first, to replicate the findings of the *Quest* experiments using different materials, and, second, to test whether coherence effects would be present in tasks that hinge entirely on probabilistic evidence, including evidence expressed in numeric terms. The experimental design enabled a comparison between the descriptive power of coherence-based reasoning and Bayes Theorem, considered by many to be the reigning conceptual framework for the understanding of evidence law. Additionally, we sought to increase the external validity of coherence-based reasoning by using a larger and more diverse pool of participants. In this study, participants, all of whom were eligible for jury service, were presented with a task that would be demanded of actual jurors.⁵⁰

As in the *Quest* experiments, participants were first presented with, and asked to draw factual inferences from, a pretest that consisted of seven apparently isolated vignettes. For example, a vignette that concerned an eyewitness identification of a person was followed by a factual question about the likelihood that the identification was correct.⁵¹

The second phase of the experiment presented participants with a whodunit case that involved a theft of money from a construction company's safe. The sole question in the case concerned the identity of the perpetrator. One of the company's employees, Jason Wells, was suspected of committing the theft. Participants were asked to play the role of jurors assigned to decide the case.⁵²

The pieces of evidence—all circumstantial—were unrelated to each other. Four items tended to incriminate the defendant: an eyewitness identification; a security camera that captured a car like the one defendant drives screeching out of the parking lot soon after the crime was committed; a call the defendant made on his cell phone from the vicinity of the building soon after the crime; and an apparent grudge that he held against the company (these items will be labeled “inculping items”).⁵³ Two of the inculping items were defined in

the website “About.com.” As an incentive, participants entered a lottery for \$200, with odds of winning at 1/200. Participation was limited to American citizens over the age of eighteen.

⁵⁰ This extension was also intended to rule out the remote possibility that the previous findings were driven by what the participants might have considered to be an experimental demand. It is conceivable that the participants in the *Quest* experiments imposed coherence because they believed that coherence is the way judges (whom they were simulating) perform the task. There is no reason to suspect that they hold the same conception of juror decision making.

⁵¹ The factual questions were followed by questions pertaining to participants' general beliefs concerning the respective facts. For a discussion on the findings regarding the general beliefs, see Part I.B.4.

⁵² In some of the studies, participants were asked to play the role of an arbitrator in a disciplinary hearing brought by the company against Jason Wells.

⁵³ For example, the evidence pertaining to the eyewitness identification read as follows:

The night of the crime, a technician was called in to repair the photocopying machine in Big

numeric values.⁵⁴ Three pieces of evidence had an exonerating effect: he was seen far away from the crime scene later that evening; payments he made after the crime came from legitimate family transactions; and it was possible that he was working harder to make amends with the company (these items will be labeled “exculpating items”). The evidence was virtually identical to the factual issues contained in the vignettes. In all, the evidence was sufficiently multifarious and balanced so as to create a complex case.

Participants were next presented with the arguments made by the attorneys of both sides of the case. The arguments pertained to inferences from the underlying evidence. Participants were asked to render a verdict and to rate their confidence in the decision.⁵⁵ They were then asked to rate their agreement with the inferences as argued by the parties. The questions were phrased in terms of the likelihood of the defendant’s guilt, given each piece of evidence, a phrasing that approximates Bayesian judgments.⁵⁶ These questions were essentially identical to those asked on the pretest.⁵⁷

Consistent with our predictions and with the findings in the *Quest* case, the ratings of the facts shifted considerably and consistently toward coherence with the eventual verdict. As seen in Figure 4, at the vignette phase, the facts are ambiguous and nonprobative: there is no significant difference between the mean pretest ratings of the inculcating and exculpating items, and there are no differences between the mean ratings given by participants who ultimately decided against the defendant and ratings given by those who acquitted him. At the

Buildings’ office. The technician testified that as he was on his way out of the office, he saw a person rushing out of the bookkeeper’s office and then disappearing down the stairs. The time then was about 7:15 PM. The next day, police detectives asked the technician to go to identify the defendant. When they got to Jason’s office, the technician said that he recognized Jason as the man he saw the night before leaving the bookkeeper’s office. When asked how certain he was, the technician responded that he was completely certain that it was Jason. He added that he had seen Jason in the building once or twice before.

⁵⁴ One numerically defined piece of evidence was the uniqueness of the match with the defendant’s car. A video camera depicted a car screaming away from the scene of the crime, but its license plate could not be deciphered. The car make and color matched the defendant’s car, and he was seen driving it to work that day. Participants were told that only 1 percent of the cars in the county matched the description. Another piece of evidence concerned a phone call Jason made just minutes after the crime. The cell phone company identified the location of the call as very close to his office. This information undermined Jason’s alibi (of being at home). Participants were also told that the location of calls was accurate in 98 percent of cases.

⁵⁵ To obtain high sensitivity in responses, the confidence rating was elicited on an eleven-point scale, ranging from 1 (low), through 6 (medium), to 11 (high).

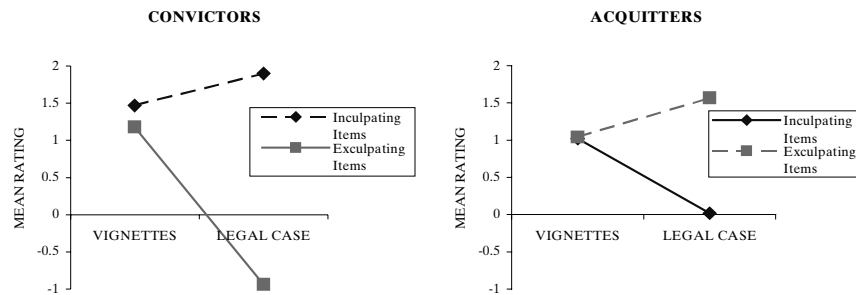
⁵⁶ For example, participants were asked to rate their agreement with the factual statement: “The technician’s identification of Jason makes it likely that the person hurrying out of the bookkeeper’s office was in fact Jason.”

⁵⁷ Responses to the inference questions were collected on the same eleven-point response scale.

second measurement, in contrast, we see large differences between participants who found the defendant guilty and those who found him innocent. For the former, we observe an increase in ratings of the inculpatory items and an acute decrease in that of the exculpatory items; opposite shifts are seen for those who decided in his favor.⁵⁸ In other words, by the point of decision, the mental models of the evidence are skewed toward strong support of the respective verdicts. Notably, these coherent representations were not inherent to the perceptions of the evidence itself. Rather, they resulted from the polarizing transformation of the originally ambiguous and nonprobative evidence toward either one of the two coherent mental models, each of which was decidedly probative with regard to the respective verdict.⁵⁹

⁵⁸ Figure 4 presents the mean ratings separately for participants who decided against the defendant (“Convictors,” in the left panel) and participants who decided in favor of the defendant (“Acquitters,” right panel). The graphs represent the mean rating for inculpatory items and exculpatory items separately.

FIGURE 4



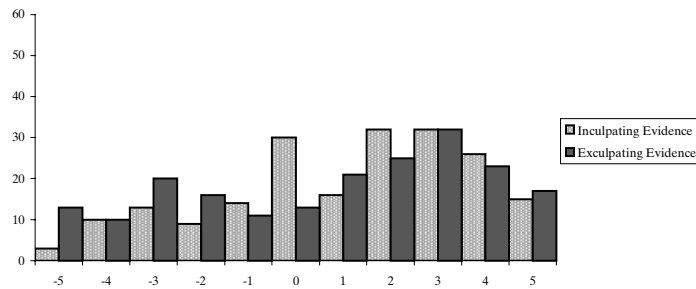
As seen on the left panel, for participants who convicted the defendant, ratings of inculpatory items increased relative to the ratings of virtually the same evidence at the vignette phase (from a mean of 1.47 to 1.90), while the ratings of the exculpatory items dropped dramatically (from 1.18 to -0.94). Opposite shifts are found for those who acquitted the defendant, with ratings of inculpatory items decreasing relative to the vignette phase (from a mean of 1.02 to 0.02), and ratings of the exculpatory items increasing (from 1.04 to 1.57).

⁵⁹ Another way to depict the polarization of the perceptions of the evidence is a histogram that plots the aggregate number of ratings given to the inculpatory items (light-colored bars) and exculpatory items (dark-colored bars). Figure 5 presents the ratings only for the participants who convicted the defendant. The horizontal axis represents the scale of ratings, ranging from -5 (“strongly disagree”) to +5 (“strongly agree”). The vertical axis represents the frequency of those ratings, that is, the number of times each rating was given by this population. At the vignette stage (top panel), the ratings of both inculpatory and exculpatory items are distributed relatively evenly across the scale, suggesting that these participants found both types of evidence to be similarly true. At the decision phase (lower panel), there is a polarization in the ratings, with more positive ratings given to inculpatory evidence and more negative ratings given to exculpatory evidence. This is, of course, another manifestation of the coherence shifts. Similar polarization (not depicted here) was found among participants who acquitted the defendant.

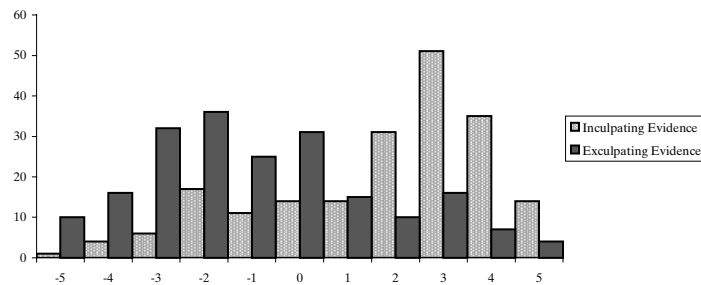
As predicted, and consistent with the *Quest* findings, participants demonstrated remarkably high levels of confidence in their decisions, regardless of their verdicts. Virtually none of the subjects were neutral or uncertain about a verdict: only 11 percent of the participants gave confidence ratings values of between 1 and 5 (on a scale ranging from 1 to 11), and 49 percent gave ratings of 9 and above. We next sought to test for a relationship between participants' ratings of the evidence and their reported levels of confidence. We found that the confidence levels were statistically related to the difference between the ratings of the inculpatory and exculpatory evidence, so that higher discrepancies in perceptions of evidence resulted in higher levels of confidence.⁶⁰ Finally, levels of confidence were also positively related to the magnitude of the coherence shift, so that participants who displayed higher differences between the ratings of the evidence at the two points of measurement also reported higher levels of confidence.⁶¹ In

FIGURE 5

VIGNETTES



LEGAL DECISION



⁶⁰ For this analysis, we split the participants (regardless of their decisions) into groups of high confidence (ratings of 9 and above), and low confidence (ratings of 8 and below). The discrepancy between the inculpatory and exculpatory difference was found to be 1.90 for the low confidence participants and 3.00 for the participants who reported high confidence. This difference was highly significant.

⁶¹ Using the same split between high and low levels of confidence, we compared the discrepancy between the inculpatory and exculpatory ratings at the two points of measurement. The

other words, the greater the transformations of the mental models, the higher the confidence experienced.

2. The lack of awareness of coherence shifts.⁶²

The second feature of coherence-based reasoning concerns its metacognitive dimension—the degree to which people are aware of the process, particularly of their shifting perceptions of the task variables. We hypothesized that coherence shifts transpire without awareness—that they are mostly an automatic, rather than a controlled and conscious, form of cognitive processing.⁶³

Testing for metacognition typically requires indirect measures. Our method was based on the assumption that one must have knowledge of both the original and the subsequent ratings of the task variables to be aware of coherence shifts. A failure to recall the original state precludes the possibility of comparing the two states, and thus of noticing change. We ran two studies that were identical to the first *Quest* study, except that we added another instrument. After the decision task was completed, the participants were asked to recall the ratings they gave at the vignette phase.⁶⁴

Results showed that the participants recalled their original ratings inaccurately: the ratings they reported actually approximated their current, post-coherence shift ratings. This finding suggests that people's awareness of coherence shifts is limited; people feel that their current beliefs are the ones they held all along. As discussed below, this phenomenological state maintains the aura of rationality and legitimacy in the decision-making process.

discrepancy was 1.23 for the low confidence participants and 1.71 for participants who reported higher confidence. This difference was highly significant.

⁶² These studies were published as Experiments 2 and 3 of Holyoak and Simon, 128 *J Exper Psych: Gen 3* (cited in note 5).

⁶³ Mental processes are generally said to be automatic when they occur outside of awareness, require no conscious cue for initiation, consume little cognitive effort, and are readily stoppable. All four criteria do not have to be present in order to treat a process as automatic. On the increasingly important distinction between automatic and controlled processing, see Daniel M. Wegner and John A. Bargh, *Control and Automaticity in Social Life*, in Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, eds, 1 *The Handbook of Social Psychology* 446 (McGraw-Hill 1998); John A. Bargh, *The Four Horsemen of Automaticity: Awareness, Intention, Efficiency, and Control in Social Cognition*, in Robert S. Wyer, Jr. and Thomas K. Srull, eds, 1 *Handbook of Social Cognition* 1 (Lawrence Erlbaum 2d ed 1994).

⁶⁴ Participants had not been told of this final recall task in advance.

3. The facilitative function of coherence shifts: not mere post hoc justification.⁶⁵

The next series of experiments examined the important theoretical question of the role that coherence shifts play in the decision-making process. It could be argued that the shifts play no active role in the process itself, but serve only as post hoc rationalizations for decisions driven by other factors or different mechanisms. This is the view offered by cognitive dissonance theory, which posits that attitudes and preferences change exclusively due to post-decision regret. Only after a person commits herself to a course of action does regret cause the arousal of dissonance, which is then reduced by rationalizing the decision.⁶⁶ We sought to test this aspect of coherence-based reasoning.

In a series of studies based on the *Quest* case, we induced participants to delay their decisions and measured their preliminary rating of the arguments during the period of delay.⁶⁷ Contrary to dissonance

⁶⁵ These studies were reported as Experiment 1 of Holyoak and Simon, 128 *J Exper Psych: Gen 3* (cited in note 5), and as Experiments 1–3 of Simon, et al, 27 *J Exper Psych: Learning, Memory & Cognition* 1250 (cited in note 5).

⁶⁶ See Leon Festinger, *Conflict, Decision, and Dissonance* 30–31, 153 (Stanford 1964); Leon Festinger, *A Theory of Cognitive Dissonance* 34–35, 42–47 (Stanford 1957). Jerome Bruner criticized dissonance theory, claiming that “the most interesting aspects of cognition are those that precede the making of decisions rather than those that follow [it].” Jerome Bruner, *Discussion*, in Jerome S. Bruner, et al, *Contemporary Approaches to Cognition: A Symposium Held at the University of Colorado* 152 (Harvard 1957). Bruner commented that dissonance theory amounted to a “rather autistic tradition.” *Id.* Coherence-based reasoning can be seen as the experimental corollary of Bruner’s observation.

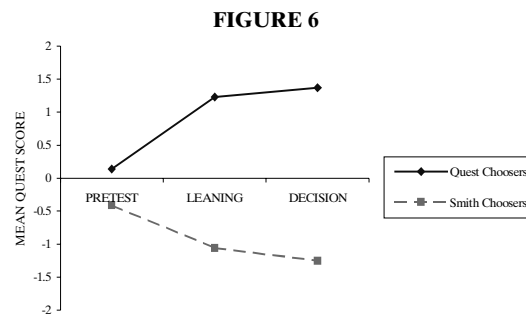
For reviews of the relationship between cognitive dissonance theory and Gestalt psychology, see Read, Vanman, and Miller, 1 *Personality & Soc Psych Rev* at 44–46 (cited in note 28) (discussing the reliance of cognitive dissonance on Gestalt psychology); Hazel Markus and R.B. Zajonc, *The Cognitive Perspective in Social Psychology*, in Gardner Lindzey and Elliot Aronson, eds, 1 *Handbook of Social Psychology* 137, 197–214 (Random House 3d ed 1985). For a critical view of cognitive dissonance as a field of decision making, see Simon and Holyoak, 6 *Personality & Soc Psych Rev* at 283–85 (cited in note 5).

⁶⁷ In these studies, participants completed essentially the same procedure as did those in the basic condition, except that in the initial instructions participants were told that before reaching a verdict they should wait to hear the verdict of another judge in a very similar case, as the other verdict would provide additional information relevant to their decision. We also explained that the other judge was more senior and had an outstanding reputation, and that reaching different results would cause some legal confusion and be a personal embarrassment for the participants. In the meantime, they were instructed to read the case and think about it. Before receiving any new information, these participants were asked to state their “preliminary leaning” toward either *Quest* or *Smith*, and to report their evaluations of the arguments. After their response forms were collected, participants were then told that the other judge was not going to deliver a verdict after all and that they should proceed to reach a final verdict by themselves based on the facts and arguments they had read. After stating their verdict, they completed the second assessment instrument again (with a different random order of the arguments). See Holyoak and Simon, 128 *J Exper Psych: Gen* at 5–6 (cited in note 5) (Experiment 1, three-phase participants only); Simon, et al, 27 *J Exper Psych: Learning, Memory & Cognition* at 1252–53 (cited in note 5) (Experiment 1).

theory, significant coherence shifts were observed at this early stage.⁶⁸ This finding suggests that coherence shifts play a functional role in the decision-making process, in that the spreading apart of the variables creates dominance of one alternative over its rival, thereby enabling a confident decision. Similar predecisional shifts were observed in a study based on the *Jason Wells* case.⁶⁹ In addition to the shifts that preceded the decision, coherence shifts were also observed following the point of decision, though the post-decisional shifts were typically marginal.⁷⁰ Coherence shifts, the data suggest, precede decisions.⁷¹

⁶⁸ Figure 6 presents the results of Study 1 in Simon, et al, 27 J Exper Psych: Learning, Memory & Cognition at 1250 (cited in note 5). The vertical axis represents the degree of support for the vying decisions, with positive scores supporting Quest. The graphs represent the “Quest Score” at the three points of measurement, that is, the ratings of the variables that support Quest combined with a reverse score of the variables that support Smith. Positive “Quest Scores” represent strong agreement with Quest’s arguments and disagreement with Smith’s arguments.

Participants who chose the Quest verdict (top graph) shifted to high support with Quest’s arguments by the interim measurement, and opposite shifts are noticeable for participants who decided for Smith (bottom graph). A small additional shift is observed between the interim measurement and the final decision. The latter shift can be understood to have occurred post-decisionally.



⁶⁹ See Experiment 3 of the *Jason Wells* study in Simon, Snow, and Read, *Evidence Judgments* (cited in note 5). Further evidence of predecisional shifts comes from an experiment described below in which participants were found to re-shift their mental models in the middle of a decision following introduction of new evidence. See Part I.B.5. Predecisional shifts were also observed in a study (not reported here) that involved a choice between two job offers. Judgments made during the delay phase displayed substantial shifts. See Study 1 in Simon, Krawczyk, and Holyoak, *Construction of Preferences* (cited in note 5).

⁷⁰ In most of these conditions, the post-decisional shifts did not reach statistical significance. See Holyoak and Simon, 128 J Exper Psych: Gen at 7, 10–11, 14 (cited in note 5); Simon, et al, 27 J Exper Psych: Learning, Memory & Cognition at 1253–56 (cited in note 5).

⁷¹ It is theoretically possible to interpret the findings of coherence shifts at the delay phase not as predecisional shifts. Participants could be deemed to have already made a decision by that point, so that the observed shifts are effectively post-decisional. To refute this possibility, we ran two studies in which the *Quest* case was presented not as a decision task at all. For example, in one study, participants were not asked to play the role of the judge nor were they asked to make any decision; rather, they were instructed to memorize the arguments made by the lawyers in preparation of a memory test. Simon, et al, 27 J Exper Psych: Learning, Memory & Cognition 1250 (cited in note 5) (Experiment 2). We found substantial coherence shifts at this point. This finding suggests that coherence shifts are not post-decisional phenomena. They are better understood as a general mechanism of cognition in conditions of complexity. For a similar experiment,

4. The depth and breadth of coherence effects.⁷²

All mental processing draws closely from one's background knowledge.⁷³ A decision to cross a street, for example, is contingent on one's experience-born knowledge about vehicles, motion, and driver behavior. A choice to form a friendship is influenced by one's knowledge of cues for trustworthiness, love, selfishness, and the like. Thus, when a juror in a criminal case is presented with testimony about an eyewitness identification, his judgment of that evidence is bound to be informed by his background beliefs about the accuracy of eyewitness identification in general. There is a natural coherence between specific inferences and the decision-maker's background belief system.

This relationship, however, is hardly fixed. Take, for instance, a juror who generally believes that people often misidentify strangers and who thus tends to treat a certain eyewitness identification with suspicion. Despite his doubt, the juror might be convinced by some other piece of inculpatory evidence—say, the fact that a car seen screeching away from the crime scene is the same car that the defendant drives. If the latter piece of evidence is sufficiently compelling, this juror would lean toward conviction, and as a result of coherence effects, he would also become more likely to endorse the eyewitness evidence. The question, then, is what happens to the relationship between the judgment of the particular eyewitness testimony and the now contradictory belief about the accuracy of eyewitness identification in general.

One possibility is that an inconsistency arises, presumably resulting in discord between the shifting inference and the background belief system. Such discord could potentially have a destabilizing effect on the emerging conclusion, resulting in lower confidence. Another possibility is that the background beliefs shift too, and thus maintain coherence with the respective inferences. For example, as the juror leans toward conviction and increasingly believes the prosecution's eyewitness, his belief about eyewitness accuracy in general becomes less skeptical. If this were the case, it would mean that coherence is a pervasive and powerful feature that spreads beyond the specific task variables and into background knowledge structures.

Experiments using the *Jason Wells* materials indicated that the latter possibility is correct. Six questions about participants' general

see id (Experiment 3). For a detailed discussion of this issue, see Simon, Snow, and Read, *Evidence Judgments* (cited in note 5).

⁷² The results concerning the depth of coherence effects were obtained in all of the *Jason Wells* experiments reported in Simon, Snow, and Read, *Evidence Judgments* (cited in note 5). The results of the breadth of coherence effects were reported in Experiment 3 of Holyoak and Simon, 128 *J Exper Psych: Gen* 3 (cited in note 5).

⁷³ See note 30 and accompanying text.

beliefs were included at both the vignette phase and the final phase,⁷⁴ and these beliefs were also referred to in closing arguments. We compared the ratings of the same questions at both phases and found that the background beliefs shifted significantly toward coherence with the respective verdict,⁷⁵ though, not as strongly as the shifts observed in the factual inferences. This difference is not surprising since background beliefs are presumably learned and rehearsed, and therefore more stable than specific judgments in ad hoc tasks.

It is noteworthy that coherence effects not only run deep, they also run wide; they can spread from one case to the next. In another experiment, after deciding the *Quest* case, participants were asked to decide a second, unrelated case that shared a single common issue.⁷⁶ We found that coherence spread from the first verdict, through the common issue to the verdict of the second case, and reached as far as causing shifts in other variables in that case.

5. The effects of indirect influences.⁷⁷

Another feature of coherence-based reasoning that warrants attention concerns the effect of changes in one task variable on other variables. Connectionist theories posit that any variable can potentially influence the entire network, including variables to which it is not directly linked. In a study based on the *Quest* case, one-half of the participants received an account in which the defendant Smith was described as a benevolent person, and the other half received a description of a malevolent defendant. We expected that this manipula-

⁷⁴ For example, following a question about the accuracy of a specific eyewitness identification, participants were asked to report their general belief by rating their agreement with the following sentence: "In general, when people identify someone whom they've seen once or twice before the identifications are accurate." The same question was used at both measurement phases.

⁷⁵ For participants who convicted the defendant, ratings of beliefs associated with innocence declined (from 0.47 to -0.17); and for those who acquitted him, the ratings of inculpatory items decreased (from a mean of 1.26 to 0.37), while ratings of the exculpatory items increased (from 0.78 to 1.11). For participants who decided to convict, inculpatory evidence decreased slightly relative to the ratings at the vignette phase (from a mean of 1.87 to 1.69). This mild decrease was the only shift that did not match predictions. All other shifts occurred as predicted, and the interactions were highly statistically significant.

⁷⁶ The second case involved a contract dispute over the size of the bonus to be paid to the employees of an Internet company. The contract contained a clause that was argued to hinge on a comparison of the Internet company to either the newspaper or the local phone company—the same analogy involved in the *Quest* case. This study is published as Experiment 3 in Holyoak and Simon, 128 *J Exper Psych: Gen 3* (cited in note 5).

⁷⁷ The results for the *Quest* study discussed in this Part, based on the participation of eighty students, are published as Experiment 3 in Holyoak and Simon, 128 *J Exper Psych: Gen 3* (cited in note 5). The results for the *Jason Wells* case are reported in Studies 1 and 4 of Simon, Snow, and Read, *Evidence Judgments* (cited in note 5). Two hundred and eighty-six people participated in this study via the Internet. They were recruited as in the previous studies.

tion would affect the verdicts so that more verdicts favoring the defendant would be rendered in the former case than in the latter. The crucial question, however, was whether and how this manipulation would influence the ratings of the other variables involved in the case, none of which had any plausible relation to Smith's character.

As expected, the information about Smith's previous conduct had a strong influence on the distribution of the verdicts. Of the participants who received positive information about Smith, 72 percent rendered verdicts in his favor, whereas only 22 percent of those who received negative information decided for him. More importantly, we analyzed the ratings of all of the other variables in the case to see whether they were influenced by the new information. All the variables shifted to cohere with the chosen decision.⁷⁸ In other words, the manipulated piece of information had a substantial effect on variables with which it had no plausible relationship. For example, manipulating the defendant's character changed participants' views on the analogy likening the Internet to a newspaper and on the appropriateness of regulating free speech over the Internet. Similar findings of indirect influences were made in a study based on the *Jason Wells* case. The introduction of DNA evidence that inculpated or exculpated the defendant resulted in predicted changes in all the other pieces of evidence.⁷⁹

Another experiment based on the *Jason Wells* case was designed to get a closer look at the dynamics of coherence-based reasoning by examining what happens when people change their minds during the decision-making process. In light of the finding that coherence effects occur before decisions are made,⁸⁰ one might believe that people should rarely change their minds in response to new information and that when they do, they are left in a state of strong incoherence with their previous mental models. Yet, experience suggests that, under some conditions, people respond to new evidence that contradicts

⁷⁸ Combined mean shifts were from -0.27 to $+0.93$, and from $+0.01$ to -0.71 . We omitted assessment of the defendant's motive, one of the original six variables, due to the possibility that this variable would be directly affected by the information about Smith.

⁷⁹ All participants were presented with the same basic evidence, with one key exception. One-half of the participants were presented with evidence of a positive DNA match between the defendant and the perpetrator, whereas the other half was presented with evidence of a negative DNA match. As expected, the DNA manipulation affected the distribution of verdicts, with 32 percent voting to convict with a negative match and 69 percent voting to convict with a positive match. Again, the manipulation of a single piece of unrelated evidence altered the entire factual pattern. See Experiment 1 in Simon, Snow, and Read, *Evidence Judgments* (cited in note 5).

Similar indirect influences were also observed in a job choice experiment (not reported here), where we found that manipulating the favorability of the office location (located adjacent to a fun shopping center versus a dreary industrial area) affected both the job choice itself and also the participants' preferences for other variables (for example, salary and length of commute). See Simon, Krawczyk, and Holyoak, *Construction of Preferences* (cited in note 5).

⁸⁰ See Part I.B.3.

prior leanings by changing their minds,⁸¹ and that changing one's mind does not seem to have notable influences on the resulting decisions.

The experiment tracked the representation of the evidence after we introduced strong new evidence that was inconsistent with the person's previous leanings. One possibility is that since the preexisting evidence cohered with the previous decision, the new decision would not be supported by a coherent set of variables. The alternative possibility is that a second shift would occur—in the opposite direction—ultimately leading to a newly formed state of coherence. Prior to the presentation of the evidence, participants were told that an important piece of evidence was expected to become available at a later stage, and that they were therefore advised to withhold deciding the verdict until they received the missing evidence. During the delay, participants were asked to indicate their “tentative leanings,” and to rate the evidence already presented.⁸² Following the delay, one-third of the participants were presented with incriminating evidence, one-third with exonerating evidence, and one-third were told that no new information would be made available.⁸³ Participants were then asked to decide the verdict, report their confidence, and rate the evidence again.

The results showed that a majority of participants followed their initial leanings, but some did switch their verdicts.⁸⁴ Not surprisingly, participants who did not switch their verdicts displayed coherence shifts that resembled those observed in the previous delay studies—that is, a significant shift from the pretest to the interim judgments, and then a further, more moderate shift at the decision phase.⁸⁵ Of particular interest, though, were those participants who switched their

⁸¹ This is not to say that people always respond to the new evidence or that the responses are always adequate. Indeed, in some contexts, people respond very poorly. See, for example, research on confirmation bias, Nickerson, 2 *Rev Gen Psych* at 175 (cited in note 12); and groupthink, Irving L. Janis, *Groupthink: Psychological Studies of Policy Decisions and Fiascoes* 2–13 (Houghton Mifflin 1982).

⁸² Participants were reassured that any assessments they would make at this point would not be binding and that they would later “be free to make whichever decision you think is appropriate.”

⁸³ The new evidence was based on information obtained from the cell phone company as to the defendant's location soon after the money was stolen. In the incriminating condition, testimony was presented that the defendant had made a call on his cell phone from several blocks away from his office just minutes after the crime had occurred. In the exonerating condition, evidence was presented that the defendant had made a call on his cell phone from a location near his house, which was about forty-five minutes from the scene of the crime. In the third condition, participants were told that no new evidence would be obtained and they were advised to decide the case with the available evidence.

⁸⁴ About one of every five participants “switched.” Of the participants who originally leaned toward acquittal, 165 subjects were “consistent acquitters,” and 23 “switched” to conviction. Of the 121 who initially leaned toward conviction, 86 maintained a consistent preference for conviction, whereas 35 “switched” to acquittal.

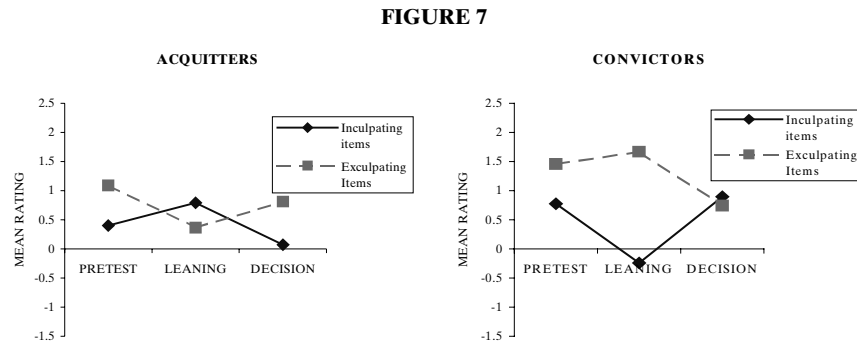
⁸⁵ See note 68 (illustrating predecisional shifts).

verdicts. Switchers displayed a unique pattern; their ratings dovetailed with their new verdict preferences. Their ratings first shifted from pretest levels to a state of coherence with their initial leanings, but as their preferred verdicts changed, so did their ratings of the evidence.⁸⁶ The ratings of the evidence, then, shifted twice—in opposite directions—from an initial state of complexity toward coherence with whichever verdict seemed more appealing at that time. Interestingly, switchers were no less confident than those who did not switch.

The findings of this tracking experiment also support the conclusion that coherence shifts occur before participants commit themselves to verdicts.⁸⁷ Participants showed strong coherence shifts at the interim measurement, despite the unavailability of supposedly important evidence, and despite being urged to withhold the decision.

In all, the experiments showing the effects of indirect influences provide the additional methodological benefit of strengthening the basic finding of coherence shifts. In these experiments, we showed not only that people drift spontaneously toward polarized mental models, but that they can be driven to do so by experimental manipulation.⁸⁸

⁸⁶ Figure 7 shows the ratings of the evidence for Switchers.



The left panel corresponds to participants who initially tended toward conviction but subsequently voted to acquit, whereas the right side panel corresponds to those who initially tended toward acquittal but ultimately convicted. Note, for example, that participants who ultimately acquitted (left panel) first reported higher evaluations for the guilt items and lower evaluations for the innocence items, but then switched to lower evaluations for the inculpatory items and higher evaluations for the exculpatory items. The opposite trend is observed for participants who switched from innocence to guilt (right panel).

⁸⁷ See Part I.B.3.

⁸⁸ It could be argued that the coherence shifts that resulted from the DNA manipulation do not provide conclusive proof of bidirectional reasoning. Unlike the *Quest* case, where the character of the defendant cannot conceivably be related to the other issues, such as the analogy of the Internet to a newspaper or a telephone system, in identity cases, the inferences could theoretically affect one another. For example, it could be said that a positive DNA match warrants a strong belief of the prosecution's eyewitness because the defendant was proven to have been at the crime scene, and that justifies treating his alibi claim with suspicion. This possibility cannot be ruled out, though I doubt that these types of inferences can fully explain the strong coherence

6. The effect of predilections—motivations and attitudes.

One of the most important developments in experimental psychology over the past two decades has been the retreat from the view that cognitive psychology can be reduced to pure models of information processing. The prevailing consensus is that motivations guide and interact with the cognitive system, influencing how people reason and perceive their environments.⁸⁹

We sought to test whether motivations interact with coherence effects.⁹⁰ A series of experiments revealing shifts in estimates of the likelihood of winning a horse race provide partial empirical support. We found that as people approach the decision of placing a gamble on a horse, they report increasingly higher estimations of winning. As we interpret this finding, the positive valence of winning the race influences the mental model of the race variables so that they cohere with the desired outcome, thus causing an increase in the purportedly objective criterion for the choice—the perceived odds of winning.⁹¹

Stronger empirical support has recently been collected in a study that offers a more direct test for the relationship between motivated outcomes and coherence-based reasoning. In this case, called *Dixon v Providential Life*, a widow is suing an insurance company for the proceeds of her late husband's life insurance.⁹² The company refused compensation based on its view that the death was a suicide. The facts are ambiguous with respect to the cause of death. Participants were assigned to play the role of the parties' lawyers. The findings show strong coherence effects supporting the desired conclusions: partici-

shifts observed across all of the items. For example, it is not so easy to conjecture a rational link between the DNA evidence and the inference that the defendant's grudge against his employer makes it more likely that he was the culprit.

⁸⁹ For discussions of the rapprochement, see Kunda, *Social Cognition* at 211–63 (cited in note 14); Eliot R. Smith, *Social Cognition Contributions to Attribution Theory and Research*, in Patricia G. Devine, David L. Hamilton, and Thomas M. Ostrom, eds, *Social Cognition: Impact on Social Psychology* 77, 82 (Academic 1994) (concluding that “cognition and motivation are inseparable”); Richard M. Sorrentino and E. Tory Higgins, eds, *Handbook of Motivation and Cognition* 3–19 (Guilford 1986).

⁹⁰ Theoretical support for this proposition can be found in Paul Thagard's conception of emotional coherence. Thagard argues that coherence systems have emotional valence. Valence indicates such features as likeability, desirability, or similar kinds of positive or negative attitudes. Variables can be seen to be related to one another by valence constraints. A variable that is strongly valenced can affect the valence of the entire mental model. See Thagard, *Thought and Action* at 170–77 (cited in note 25).

⁹¹ See Brownstein, Read, and Simon, *Pre-decision Reevaluation* (cited in note 5); Brownstein, Simon, and Read, *Betting Decisions* (cited in note 5). The interpretation of the role of coherence is somewhat tentative because we have not yet tested for the ratings of the other task variables that were available to the participants.

⁹² The materials are taken from Edward R. Stein and Frank D. Rothschild, eds, *Dixon v. Providential Life Insurance Co.* (National Institute for Trial Advocacy 2000). I am grateful to NITA for permission to use these materials.

pants report optimistic predictions of the jury decision; they make settlement offers that are skewed in their clients' favor; they rate the contested facts of the case as supportive of their side; they believe that fairness is on their side; they find their own witnesses to be more credible than the opponent's and maintain that jurors would believe the same; they believe that jurors would like their client more than the opponent; and more.⁹³ We observe a strong and significant correlation between the strength of desire and the magnitude of the distorted views of the case.

We have also found evidence that supports the proposition that coherence effects interact with the decision-maker's preexisting attitudes, particularly those embedded in the person's enduring value system.⁹⁴ For example, we found that, consistent with previous research,⁹⁵ participants with strong pro-death penalty attitudes were more likely to convict the defendant than participants with anti-death penalty attitudes. As predicted, the mental models of the evidence and beliefs cohered with the corresponding verdicts, so that people with strong pro- or anti-death penalty attitudes interpreted all the evidence in the case in a way that cohered with conviction and acquittal, respectively.⁹⁶

7. The transitory nature of coherence.⁹⁷

In another recent experiment, we tested the endurance of coherence shifts; we intended to explore whether coherence shifts cause long-term changes in people's attitudes, or whether they are merely transitory states. The former hypothesis suggests that decisions bind one to certain positions for future decisions, with the result of exacerbating complexity and incoherence under different circumstances. The latter hypothesis implies that the cognitive system is flexible and adaptive, and enables representations to shift toward coherence on the task at hand without encumbering the person for future decisions.

⁹³ See Simon, Snow, and Read, *Self-Serving Bias* (cited in note 5). We are continuing to run studies intended to replicate these findings and obtain more detail on the interaction between motivation and coherence phenomena.

⁹⁴ Attitudes embedded in one's enduring values are also called ego-involving attitudes. See Blair T. Johnson and Alice H. Eagly, *Effects of Involvement on Persuasion: A Meta-analysis*, 106 *Psych Bull* 290, 305–11 (1989) (discussing the effects of involvement on attitude change).

⁹⁵ See, for example, Jane Goodman-Delahunty, Edith Greene, and Winston Hsiao, *Constructing Motive in Videotaped Killings: The Role of Jurors' Attitudes toward the Death Penalty*, 22 *L & Human Behav* 257, 265–67 (1998); William C. Thompson, et al, *Death Penalty Attitudes and Conviction Proneness: The Translation of Attitudes into Verdicts*, 8 *L & Human Behav* 95, 109–11 (1984).

⁹⁶ See Study 2 in Simon, Snow, and Read, *Evidence Judgments* (cited in note 5).

⁹⁷ These findings will be reported in Dan Simon, Daniel C. Krawczyk, and Keith J. Holyoak, *Coherence Comes and Goes: Testing the Endurance of Coherence Effects* (in progress) (on file with author). Thus far, forty-one participants have completed all three measurements discussed in this Part.

To study this question, we used materials that had elicited coherence shifts in a job choice task (an experiment that is not reviewed in detail).⁹⁸ In the current experiment, we measured the ratings three times: on the pretest, at the time of making the decision, and then again one week later.⁹⁹ As with all the other coherence experiments, the ratings of the attributes shifted from the pretest to the decision phase. The crucial test was to observe the ratings on the last measurement. We found that the coherence shifts had essentially disappeared by that point.¹⁰⁰ These results suggest that at least under some circumstances, coherence effects are transient: they peak at the point of decision, but then recede as the point of decision passes.

8. Moderating (debiasing) coherence.¹⁰¹

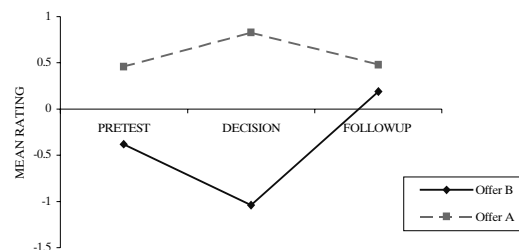
Another version of the *Quest* case examined whether coherence effects can be undone, or at least moderated. In other fields of research, attempts to “debias” people and prevent them from making errors in judgment and choice tasks have had mixed results. Direct approaches, like informing participants of the existence of the bias and imploring them to “try harder” or to “be unbiased” have been generally unsuccessful.¹⁰² Some success has been obtained with techniques

⁹⁸ See Simon, Krawczyk, and Holyoak, *Construction of Preferences* (cited in note 5). In that original study, participants’ preferences for the relative attributes of competing job offers—namely, the salary, length of commute, vacation package, and type of office—shifted toward greater support for the job offer that was eventually chosen.

⁹⁹ At the last phase, participants were not asked to make the decision again. They were asked only for their ratings of the attributes.

¹⁰⁰ Figure 8 shows the compounded values of the ratings of the attributes of the two job offers and the weights assigned to them, at three points in time: at pretest, decision, and follow-up. Note that the ratings were coherent with the decisions only at the point of making the decision. One week after the decision was made, the ratings were not different from their spontaneous state preceding the decision.

FIGURE 8



¹⁰¹ These findings are reported in Daniel C. Krawczyk, Dan Simon, and Keith J. Holyoak, *Moderating the Effects of Decision Making by Constraint Satisfaction* (in progress) (on file with author).

¹⁰² See, for example, Baruch Fischhoff, *Perceived Informativeness of Facts*, 3 *J Exper Psych: Human Perception & Performance* 349, 352–54, 356–57 (1977) (reporting the ineffectiveness of

that induce participants to actively create mental models in which they imagine alternative conclusions by urging them to consider the correctness of the opposite conclusion¹⁰³ and to note the weaknesses of their preferred conclusion.¹⁰⁴

Participants were randomly assigned to one of three conditions—a control condition in which no debiasing instruction was given, a condition in which participants were given an instruction imploring them to “be unbiased,”¹⁰⁵ and a condition in which they were instructed to “take some time to seriously consider the possibility that the opposite side has a better case.”

The results showed that the “be unbiased” instruction made little difference as compared to the control condition, but the “consider-the-opposite” instruction had a moderating effect on coherence shifts. The magnitude of the coherence shifts in the “consider-the-opposite” condition were about one-half of what they were in the control condition, whereas the shifts in the “be unbiased” condition did not differ from those in the control condition. Based on this first study, it appears that coherence shifts can be reduced substantially by means of a simple instruction. We expected also to find lower confidence levels in the “consider-the-opposite” condition, but found in fact no such difference. More studies are required to gain a better sense of the effects of the debiasing intervention.

C. Summary and Discussion of Experimental Findings

The picture that emerges thus far from the experimentation is that decision making in complex tasks is governed by coherence-based reasoning, and this process appears to enable people to solve computationally daunting tasks with great confidence. Specifically, the experimentation reveals eight features of coherence-based reasoning:

1. The basic finding of coherence shifts.

Throughout the decision-making process, mental representation of the task shifts toward a state of coherence with the emerging ver-

exhorting participants to “work harder” or telling them about the “knew-it-all-along” bias).

¹⁰³ See, for example, Charles G. Lord, Mark R. Lepper, and Elizabeth Preston, *Considering the Opposite: A Corrective Strategy for Social Judgment*, 47 *J Personality & Soc Psych* 1231 (1984) (debiasing assimilation biases); Paul Slovic and Baruch Fischhoff, *On the Psychology of Experimental Surprises*, 3 *J Exper Psych: Human Perception & Performance* 544 (1977) (reducing the effect of hindsight bias in judgments of the predictability of scientific results).

¹⁰⁴ See, for example, Linda Babcock, George Loewenstein, and Samuel Issacharoff, *Creating Convergence: Debiasing Biased Litigants*, 22 *L & Soc Inquiry* 913 (1997) (debiasing in the context of self-serving biases).

¹⁰⁵ The “be unbiased” instructions read in part: “We would like you to be as objective and unbiased as possible in making your decision about the verdict.”

dict. Ambiguous, equivocal, and conflicting variables are transformed into coherent models, that is, lopsided and exaggerated mental representations in which the variables that support the emerging decision are strongly accepted while those that support the losing decision are dismissed, rejected, or ignored.

2. Lack of awareness.

People tend not to appreciate the incompatibility between their initial and eventual mental models of the task. This finding supports the view that coherence shifts are governed by automatic cognitive mechanisms that operate under the level of conscious awareness. Coherence shifts are part of the many important cognitive processes that take place in the background of our conscious arena, without which making sense of the world would be difficult, if at all possible.¹⁰⁶ This finding is consistent with a robust body of literature demonstrating that people have a general tendency to perceive objectivity in their reasoning processes,¹⁰⁷ and to view the world through a perspective of “naïve realism.”¹⁰⁸ The lack of awareness gives the decision-maker a false sense of constancy in his own perception of the case. The ensuing decision is thus experienced as rationally warranted by the inherent values of the variables, rather than by an inflated perception imposed by the cognitive system. The lack of awareness then helps the decision-maker maintain the phenomenological experience of rationality and objectivity.

Coherence shifts, consequently, do not represent conscious, strategic, or deceitful conduct on the part of a decision-maker; rather, they

¹⁰⁶ On the breadth and importance of automatic processing, see John A. Bargh and Tanya L. Chartrand, *The Unbearable Automaticity of Being*, 54 *Am Psychologist* 462 (1999); Wegner and Bargh, *Control and Automaticity* at 446 (cited in note 63).

¹⁰⁷ It is the coupling of biased processes with lack of awareness that makes for what has been called mental contamination. See Timothy D. Wilson and Nancy Brekke, *Mental Contamination and Mental Correction: Unwanted Influences on Judgments and Evaluations*, 116 *Psych Bull* 117 (1994). See also Nickerson, 2 *Rev Gen Psych* at 175 (cited in note 12).

¹⁰⁸ Naïve realism refers to one’s

unshakable conviction that he or she is somehow privy to an invariant, knowable, objective reality—a reality that others will also perceive faithfully, provided that they are reasonable and rational, a reality that others are apt to misperceive only to the extent that they (in contrast to oneself) view the world through a prism of self-interest, ideological bias, or personal perversity.

Robert J. Robinson, et al, *Actual versus Assumed Differences in Construal: “Naive Realism” in Intergroup Perception and Conflict*, 68 *J Personality & Soc Psych* 404, 405 (1995). See also Lee Ross and Andrew Ward, *Naive Realism in Everyday Life: Implications for Social Conflict and Misunderstanding*, in Edward S. Reed, Elliot Turiel, and Terrance Brown, eds, *Values and Knowledge* 103, 110–11 (Lawrence Erlbaum 1996). The term naïve realism is a technical term used in social psychology and has no connection to legal realism.

are the natural consequence of the normal mechanisms of cognitive processing.

3. The facilitative function of coherence shifts.

Coherence shifts occur, for the most part, before the point of the decision. This finding refutes the view, espoused by cognitive dissonance theory, that the cognitive changes serve only to justify decisions by reducing post-decisional regret. It suggests, rather, that coherence shifts play an operative role in the decision-making process—the spreading apart of the considerations into one strong and one weak set facilitates confident choice. Any observed post-decisional shifts tended to be much weaker than the predecisional ones, suggesting that the regret-induced post-decisional distortions are secondary.¹⁰⁹

4. The breadth and depth of coherence effects.

Coherence shifts occurred both in the variables directly involved in the task and in related background beliefs. For example, changes were observed not only in the rating of the probativeness of an eyewitness's testimony but also in beliefs regarding the accuracy of eyewitness testimony in general. These findings demonstrate that coherence effects are sufficiently strong so as to pervade people's belief systems and skew background knowledge toward coherence with the case at hand.¹¹⁰ Coherence shifts have considerable breadth as well. In a separate experiment, we found that the verdict in one case spilled over into a subsequent case that shared one common variable, but was otherwise unrelated.

¹⁰⁹ As I have discussed elsewhere, there are at least three additional reasons why cognitive dissonance theory fails to provide a theoretical account for complex decision making: (1) dissonance theory is limited only to single pairs of variables, whereas complex decisions contain large numbers of variables; (2) dissonance is limited only to direct, obverse relations between elements, whereas complex cases entail structural inconsistencies that are not directly opposite of one another; (3) dissonance research has also focused almost exclusively on situations that threaten the self-concept, a feature that is only marginally implicated in many domains of decision making. See Simon and Holyoak, 6 *Personality & Soc Psych Rev* at 283–84 (cited in note 5); Simon, Snow, and Read, *Evidence Judgments* at 6 (cited in note 5).

¹¹⁰ It is interesting to note that the shifting of background belief bears a structural resemblance, again, with Rawls's reflective equilibrium. In the original formulation of the reflective equilibrium, Rawls posited a two-tier mechanism consisting of moral principles and moral judgments. The weakness of the theory was in its disjunction from broader theories of justification. See Norman Daniels, *Wide Reflective Equilibrium and Theory Acceptance in Ethics*, 76 *J Phil* 256, 258–59 (1979). In his later work, Rawls added a third tier of relevant background theories. In this "wide" reflective equilibrium, moral justifications are reached by working back and forth and finding the point at which all three tiers equilibrate. See *id.* See also John Rawls, *The Independence of Moral Theory*, in Samuel Freeman, ed., *John Rawls: Collected Papers* 286, 289–90 (Harvard 1999). Here, too, one should keep in mind the difference between Rawls's prescriptive approach and the current conceptualization of coherence. See note 25.

5. The effects of indirect influences.

In a number of experiments, we found that changing one aspect of the case triggered changes throughout the mental models: not only did it alter the verdict distribution, but it also influenced the ratings of variables that were unrelated to the manipulated variable. Such circuitous and indirect influences inhere in the connectionist nature of cognition. This finding becomes particularly acute in combination with the finding that people have low awareness of the shifts. Decision-makers are likely to perceive sufficient reason to base a decision on a particular set of variables, unaware that their perception of those variables is influenced by an extraneous—possibly illegitimate—variable.

6. The effect of predilections—motivations and attitudes.

Coherence effects are not isolated from noncognitive characteristics of the task. Recently obtained findings strongly support the view that coherence-based reasoning interacts with motivations. Not only do people's choices comport with their desired goals, but the mental models of the entire task tend strongly to cohere with the respective outcome. In a separate set of studies we found that coherence effects interact with preexisting attitudes to generate mental models that cohere with those attitudes.

7. The transitory nature of coherence.

The experimental evidence illustrates that coherence can be transitory. Coherence shifts were found to decay one week after the task was completed. At least in some conditions, then, coherence can be understood as an ad hoc state that is constructed only to solve the task at hand. Dissipation of coherence enables decision-makers to approach new tasks unencumbered by previous coherence shifts. In subsequent tasks, different, perhaps opposite, pressures may activate the same variables, causing them to take on different values.

Under different circumstances, coherence effects may be more permanent. Recurring coherence shifts may leave a stronger imprint, as repeated shifts in background beliefs might operate as a form of learning. Thus, for example, a judge who repeatedly presides over criminal cases in which the rate of conviction is high (as it is in real life) might well experience increasing reinforcement of beliefs about people's guilt, and thus come to develop a guilt-prone attitudinal system. Further experimentation is required to gain better insight into the duration of coherence shifts.

8. Moderating (debiasing) coherence.

From a prescriptive point of view, an important experimental finding is that coherence shifts can be reduced by means of a straightforward moderating technique.¹¹¹ A moderating instruction to “consider the opposite” reduced coherence shifts by about one-half, so that participants’ reasons for deciding as they did were closer to their initial, pre-coherence perception of the case. The fact that the technique is simple and can be self-administered makes it a potentially powerful and expedient means of tempering the effects of coherence-based reasoning.

Undoubtedly, coherence effects have their limits. While the observed coherence shifts are substantial and highly significant from a statistical standpoint, they do not reach extreme values.¹¹² Coherence shifts are mediated by task-specific factors, most notably, by the degree of ambiguity inherent in the task. Ambiguous variables are more amenable to change;¹¹³ unambiguous variables, such as indisputable facts and paramount principles are less likely to shift.¹¹⁴ It is also likely that coherence shifts are mediated by the idiosyncratic characteristics of the decision-maker.¹¹⁵

It is also important to appreciate the exact implications of coherence effects. Coherence shifts skew the mental models and generate

¹¹¹ Since I do not propose to label coherence effects a “bias,” the intervention technique is best not referred to as “debiasing,” even though it is the same kind of intervention used in the context of some biases.

¹¹² The evaluations do not shift to the end of the scale, to values of -5 or +5. On average, the shifts were in the order of 2 units on this scale. A number of evaluations intersect the neutral line, thus changing from positive to negative values. Crossing the neutral line is notable in that participants ultimately disagree with inferences with which they previously agreed.

¹¹³ A proposition is said to be ambiguous when a multiplicity of its attributes can be isolated from one another with relative ease. Ambiguous propositions are malleable and thus especially susceptible to restructuring. For example, Christopher Hsee has shown that malleability—in his terms, elasticity—lends flexibility to the reasoning process. See Christopher K. Hsee, *Elastic Justification: How Unjustifiable Factors Influence Judgments*, 66 *Org Behav & Human Dec Processes* 122 (1996). Similarly, McGuire suggested that the greatest change is imposed on issues “most easily redefined.” William J. McGuire, *Cognitive Consistency and Attitude Change*, 60 *J Abnorm & Soc Psych* 345, 349 (1960). The malleability of ambiguous propositions has also been empirically demonstrated in Abraham Tesser and Claudia L. Cowan, *Some Attitudinal and Cognitive Consequences of Thought*, 11 *J Rsrch Personality* 216 (1977) (testing attitude polarization and finding that it increased with the level of ambiguity).

¹¹⁴ As Robert Abelson explained, cognitive restructuring runs into difficulties when it becomes “too great a distortion of reality.” Robert P. Abelson, *Modes of Resolution of Belief Dilemmas*, 3 *J Conflict Resol* 343, 345 (1959).

¹¹⁵ Research shows individual differences in people’s tolerance for inconsistency. See Donna M. Webster and Arie W. Kruglanski, *Individual Differences in Need for Cognitive Closure*, 67 *J Personality & Soc Psych* 1049 (1994); Steven L. Neuberg and Jason T. Newsom, *Personal Need for Structure: Individual Differences in the Desire for Simple Structure*, 65 *J Personality & Soc Psych* 113 (1993). Leon Festinger referred to this construct as “mental agility.” See Festinger, *Cognitive Dissonance* at 44 (cited in note 66).

an inflated sense of confidence even in close decisions, but it cannot be said that the shifts necessarily result in objectively wrong decisions, nor that they cause cardinal changes in the decision-maker's own perceptions of the case.¹¹⁶ They do, however, cause a substantial increase in the risk of error in certain circumstances. The four applications discussed in the next Part of the Article exemplify such circumstances.

II. APPLICATIONS TO LEGAL DECISION MAKING

I now bring coherence-based reasoning to bear on the legal system. This body of research offers a third view of the black box of legal decision making, and thus provides a way to break the stifling and unconstructive grip of the Rationalist and Critical approaches.

In previous writing, I applied this view to appellate judicial decision making, aiming to understand the enigmatic argumentative style that typifies legal opinions—the fact that virtually every argument cited converges to the same result. I argued that the exaggerated and lopsided portrayal of the case is not an accurate reflection of the legal materials; to a large degree, it is driven by the judge's skewed perception of them.¹¹⁷ I also argued that the resulting argumentative style has a harmful impact on the legal culture it fosters.¹¹⁸

This Article focuses on the trial. Most trials are paradigmatic examples of complex tasks. Fact-finders, typically jurors, are presented with a host of evidence—that is voluminous, fragmentary, incomplete, ambiguous, and inconsistent—from which they are expected to pro-

¹¹⁶ Cardinal shifts are described in the judgment and decision-making literature as “preference reversals.” See Paul Slovic and Sarah Lichtenstein, *Preference Reversals: A Broader Perspective*, 73 *Am Econ Rev* 596–605 (1983).

¹¹⁷ See Dan Simon, *Freedom and Constraint in Adjudication: A Look through the Lens of Cognitive Psychology*, 67 *Brooklyn L Rev* 1097, 1129–39 (2002); Dan Simon, *A Psychological Model of Judicial Decision Making*, 30 *Rutgers L J* 1, 121–41 (1998). The exaggerated argumentation is in part also the product of the practice of “padding” opinions. See Richard A. Posner, *Judges' Writing Styles (And Do They Matter?)*, 62 *U Chi L Rev* 1421, 1441 (1995). By itself, however, padding cannot fully account for the observed argumentative style. Padding can inflate the opinion by expanding and adding some arguments that were not part of the underlying decision to vote for the particular verdict, but it cannot explain how virtually every one of the arguments—including the core issues—aligns so perfectly to cohere with the verdict. Coherence-based reasoning, however, goes a long way toward explaining the phenomenon.

¹¹⁸ See Simon, 67 *Brooklyn L Rev* at 1129–37 (cited in note 117); Simon, 30 *Rutgers L J* at 130–32 (cited in note 117). I ended that discussion with an aspiration toward a different argumentative style, in which, as advocated by Richard Posner, the highest ambition of a judge faced with a difficult case would be to make “a ‘reasonable’ (practical, sensible) decision, as distinct from a demonstrably correct one.” See *id.* at 137, citing Richard A. Posner, *The Problems of Jurisprudence* 456 (Harvard 1990). I suggested then that the natural tendency to impose coherence might be overcome by means of a debiasing procedure which, as described above, has since been substantiated experimentally. Although it is rather unlikely that a formal debiasing procedure will be introduced into the working practices of the bench, the susceptibility of the cognitive process to debiasing is encouraging in that it suggests that there are prospects for informal or indirect means of reducing the excessive coherence.

duce a categorical verdict.¹¹⁹ The practices used to conduct and review trials are based primarily, though not exclusively, on the Rationalist paradigm, and thus do not always correspond to the cognitive capabilities and limitations of legal decision-makers. In this part of the Article, I present four situations in which coherence-based reasoning causes systematic errors in trials. Understanding coherence-based reasoning makes it possible to introduce interventions and procedures to overcome these breakdowns and thus provide more accurate and fair trial outcomes.

A possible objection to the following discussion is that the research examines the cognitive processes of individuals, while verdicts are decided by groups of jurors or panels of judges. Ample research shows differences between decisions made by individuals and by groups.¹²⁰ This research, however, does not demonstrate that groups are necessarily less susceptible to bias and error.¹²¹ In the context of jury verdicts, the research has shown repeatedly that group verdicts tend to correspond closely to predeliberation individual preferences,¹²² and there is little reason to believe that group deliberation purges the preceding coherence effects.¹²³

A. Correcting for Ignorance: The Timing of Jury Instructions

Jurors do not receive instructions on substantive law until late in the trial. Our experimentation showed that coherence shifts precede the making of the decision. Quite possibly, by the time jurors receive instructions, their perception of evidence has already shifted to either side of the issue, with a diminished chance of being affected by the legal instruction. This problem can be alleviated by providing jurors more extensive preliminary jury instructions.

Common law trials aspire to jury verdicts based on a thoughtful application of the correct legal rules to evidence properly admitted at trial. Hence, the trial is sequenced so that juries render their verdicts only after exposure to the evidence and arguments presented by both

¹¹⁹ For careful and detailed analyses of evidentiary complexity, see David A. Schum, *Alternative Views of Argument Construction from a Mass of Evidence*, 22 *Cardozo L Rev* 1461, 1493–1502 (2001); David A. Schum, *Evidential Foundations of Probabilistic Reasoning* 1–10 (John Wiley 1994). Schum's work follows in the tradition of Wigmore's inference networks. See John H. Wigmore, *The Problem of Proof*, 8 *Ill L Rev* 77 (1913–1914).

¹²⁰ See, for example, Norbert L. Kerr, Robert J. MacCoun, and Geoffrey P. Kramer, *Bias in Judgment: Comparing Individuals and Groups*, 103 *Psych Rev* 687, 713–15 (1996) (examining the difference between groups and individuals in their susceptibility to types of bias).

¹²¹ *Id.* at 715 (concluding that the social decision scheme model identifies “conditions under which groups are both more and less biased than individuals”).

¹²² See note 207.

¹²³ Indeed, my collaborators and I suspect that group processes amplify and compound coherence shifts, and we have proposed studies to test for these effects.

parties, and only after receiving the appropriate legal instructions from the judge. From a normative point of view, this design makes perfect sense; its viability from a cognitive perspective, however, is questionable.

The prevailing trial design rests on the assumption that the complex and vast amount of testimony, presented over the course of days and weeks, can be encoded, retained, and retrieved from memory in an unaltered state. Suspended in a state of cognitive abeyance, the juror dutifully awaits the formal announcement of the legal rules before starting to make sense of the case. Only after the rules are introduced does the juror retrieve the unadulterated evidence, sift relevant facts from irrelevant ones, and begin to generate theories, develop preferences, and lean toward a decision.

The findings from the coherence-based reasoning research are markedly inconsistent with these assumptions. Recall that in a number of experiments, we observed coherence shifts even when participants were awaiting important evidence and were instructed to delay a decision. Rather than waiting, the participants apparently processed the evidence quite vigorously and shifted toward mental models that were skewed toward either one of the verdicts.¹²⁴ The magnitude of these predecisional shifts was roughly as strong as shifts observed in full-fledged decisions.¹²⁵

These findings are consistent with research that shows a general human tendency to make sense of one's social and physical worlds proactively, even in the absence of specific processing goals.¹²⁶ Moreover, people are apt to spontaneously develop a liking or disliking toward the objects of their judgment.¹²⁷ The findings are consistent also with experiments with simulated jurors,¹²⁸ and with data obtained from

¹²⁴ See notes 67–69 and accompanying text.

¹²⁵ See notes 68–70.

¹²⁶ See, for example, Reid Hastie and Bernadette Park, *The Relationship between Memory and Judgment Depends on Whether the Judgment Task Is Memory-Based or On-Line*, 93 *Psych Rev* 258, 261–66 (1986) (arguing that much social judgment occurs on-line, rather than being based on retrospective reconstructions of features from memory); Laraine Winter and James S. Uleman, *When Are Social Judgments Made? Evidence for the Spontaneousness of Trait Inferences*, 47 *J Personality & Soc Psych* 237, 248–51 (1984) (discussing research demonstrating that the exposure to information about other people prompts spontaneous, automatic judgments of those targets, even when the explicit task requires no such judgments).

¹²⁷ See, for example, R.B. Zajonc, *Feeling and Thinking: Preferences Need No Inferences*, 35 *Am Psychologist* 151, 160–65 (1980) (finding that feelings and preferences preceded cognitive judgments).

¹²⁸ See Vicki L. Smith, *Impact of Pretrial Instruction on Jurors' Information Processing and Decision Making*, 76 *J Applied Psych* 220, 225–26 (1991) (showing a high prevalence of mid-trial verdict decisions); Saul M. Kassin and Lawrence S. Wrightsman, *On the Requirements of Proof: The Timing of Judicial Instruction and Mock Juror Verdicts*, 37 *J Personality & Soc Psych* 1877, 1883 (1979) (“To some extent, then, subjects' decisions were substantially formed very early in the trial presentation.”). Similarly, in a simulated civil trial, 34 percent and 44 percent of the two

real jurors serving in actual trials,¹²⁹ that show that preferences for verdicts—and to some degree, also final decisions—can emerge during the evidence phase of the trial.

The fact that jurors develop preferences and make up their minds before the case is submitted to them can have significant ramifications for the decision. The decision-making process is strongly influenced by the mental representation of the task variables. Incoming evidence becomes integrated into the mental model and linked to other pieces of evidence. The juror's understanding of the conceptual framework of the task influences the manner in which each piece of evidence is incorporated into the model. Ideally, jurors' frameworks will be shaped by the legal rule as instructed by the judge. The framework guides the juror as she assesses the potential significance of each piece of evidence and establishes its connections with the other evidence. For example, to appreciate testimony about a complex chain of events that led to an accident, the juror must have a correct understanding of tort liability, including rules concerning duties of care, causality, contributory negligence, and the like. To properly evaluate evidence about a criminal defendant's liability for a robbery, one ought to be familiar with the elements that constitute the crime. In the coherence experiments, there was little doubt as to how each variable fit into the conceptual framework, because the connections between the variables and the verdicts were designed to be obvious. In real life, however, people are often unfamiliar with the exact content of the legal rule. Expecting jurors to make legally correct judgments without instructing them on the law has been likened to introducing a person to the game of baseball and asking him to determine who won the game before having explained the rules.¹³⁰ Worse yet, experimental evidence

groups of participants reported that they had reached a decision at the halfway point of trial. See Martin J. Bourgeois, et al, *Nominal and Interactive Groups: Effects of Preinstruction and Deliberations on Decisions and Evidence Recall in Complex Trials*, 80 *J Applied Psych* 58, 61 (1995).

¹²⁹ A study of 1,385 jurors from 172 civil trials held in Arizona in 1995 reveals that 65 percent of the jurors developed a leaning in favor of one of the parties prior to the jury instructions, and 40 percent actually made up their minds prior to that point. See Paula L. Hannaford, et al, *The Timing of Opinion Formation by Jurors in Civil Cases: An Empirical Examination*, 67 *Tenn L Rev* 627, 628, 637, 640 (2000). A rate of 73 percent of predecisional leanings was observed in a study performed in criminal trials in 30 urban state court centers, with 3,626 jurors responding. See Paula L. Hannaford-Agor, et al, *Are Hung Juries a Problem?* 32, 63–64 (National Center for State Courts 2002). In-depth analyses conducted by the Capital Jury Project also show premature decisions with respect to the punishment in capital trials. Almost half of the 916 jurors interviewed reported that they had decided the punishment during the guilt phase of the trial. See William J. Bowers, Marla Sandys, and Benjamin D. Steiner, *Foreclosed Impartiality in Capital Sentencing: Jurors' Predispositions, Guilt-Trial Experience, and Premature Decision Making*, 83 *Cornell L Rev* 1476, 1486–88 (1998).

¹³⁰ See Willaim W. Schwarzer, *Reforming Jury Trials*, 132 *Fed R Dec* 575, 583–84 (1991) (discussing the case for giving jury instructions at the start of the trial). Anecdotally, having been introduced to the game of baseball as an adult, my first impressions of the game led me to mis-

shows that people who lack familiarity with the law tend to apply their often erroneous preconceptions about the content of the rules. Such findings have been observed with relatively familiar crimes such as burglary, kidnapping, and murder.¹³¹ There is less reason to expect accurate preconceptions when it comes to adjudicating antitrust violations or Occupational Safety and Health Agency regulations.

There is a strong reason to suspect that applying misconceived rules is bound to lead to unwarranted decisions.¹³² Coherence-based reasoning research suggests that by the time the judge divulges legal rules, jurors' mental models have already shifted, yielding strong preferences, if not complete decisions. At that point, it is more difficult to educate jurors as to the correct legal rules because the misconceived rule is already integrated into a coherent and stable mental model, and the evidence is skewed accordingly. This is not to say that jurors cannot be corrected or swayed, but that a corrective effort requires overcoming a certain degree of inertia. Recall that in the *Jason Wells* "switcher" study, after eliciting participants' tentative leanings, we gave them additional evidence, yet a majority did not switch their decision despite the strength of the newly revealed evidence.¹³³

Familiarizing jurors with the correct legal rules before they encode the evidence can increase the conformity of verdicts to those

conceive the meaning of home runs. Based on my familiarity with tennis, it seemed to me that a premium should be placed on accuracy rather than on brute force, giving the pitching team a fair opportunity to throw out the runners on the field. According to my conceptual framework of the game, home runs should have counted against the batting team.

¹³¹ Vicki Smith tested undergraduate students at Northwestern University for their knowledge of a number of familiar crimes. Most of the reported features were inconsistent with the Illinois criminal code. Thus, for example, participants defined burglary as: something of value is taken (54 percent, incorrect), occurs in home or apartment (46 percent, inaccurate), involves a break in (42 percent, incorrect), with a purpose to steal (33 percent, inaccurate). The correct elements of burglary in Illinois at the time of the research were: entering a building without authority with the intent to commit a felony. See Vicki L. Smith, *Prototypes in the Courtroom: Lay Representations of Legal Concepts*, 61 *J Personality & Soc Psych* 857, 861 (1991).

Misconceptions of criminal prohibitions were also found in research by psychologist John Darley and criminal law professor Paul Robinson. Specifically, they found that people tend to believe that the criminal code matches their own moral sense of the issues. See John M. Darley, Kevin M. Carlsmith, and Paul H. Robinson, *The Ex Ante Function of the Criminal Law*, 35 *L & Socy Rev* 165 (2001).

¹³² For experimental evidence, see Vicki L. Smith, *When Prior Knowledge and Law Collide: Helping Jurors Use the Law*, 17 *L & Human Behav* 507, 532–35 (1993); Darley, Carlsmith, and Robinson, 35 *L & Socy Rev* at 175 (cited in note 131) (illustrating the difference between people's belief of the law and the actual law).

¹³³ See note 84. It is quite possible that the experimental design contributed to the stickiness of the shifts, because asking participants for their preliminary leanings can inhibit people from changing their minds. But that does not seem sufficient to fully explain the fact that a majority of participants stuck with their preliminary leanings despite the introduction of very strong evidence to the contrary. Field data suggest that over the course of the trial, a large proportion of jurors reverse their leanings, but one-half of these changes occur before jury instructions. See Hannaford, et al, 67 *Tenn L Rev* at 638–39 (cited in note 129).

rules. Thus, I propose that jurors be provided with substantive preliminary instructions. The instructions should be determined at a pre-trial conference. The judge should explain to the jury that these preliminary instructions might well be supplemented and revised during the trial, and that the later instructions will be decisive. The judge may then add and revise the instructions during the trial to reflect changes in the plaintiff's or prosecution's case, and to prepare the jury for the defendant's case. Such updating and recapping of the instructions would not differ much from the practice of interim summations given by lawyers.

The idea of substantive preinstruction has received some attention from legal scholars and practitioners,¹³⁴ and a handful of legal psychologists have been studying the topic experimentally. The available research can assist in assessing the benefits and costs of the proposal.

The primary advantage sought is an improvement in jurors' comprehension of the law, resulting in greater conformity between the legal rules and the verdicts rendered. Based on the available evidence, the hypothesis seems to be borne out as predicted by coherence-based reasoning theory. In experiments with a complex tort case, preinstructed juries could better differentiate among the plaintiffs, so that the awarded compensation corresponded to the severity of the various plaintiffs' injuries.¹³⁵ In experiments based on criminal trials, participants who were instructed both before and after a trial applied the legal instructions to factual situations with greater accuracy.¹³⁶ In a study of actual juries that decided sixty-seven civil and criminal cases in Wisconsin, jurors who were given preliminary instructions were compared with jurors who were not preinstructed. The former group reported finding the jury instructions more helpful to the performance of their task.¹³⁷

¹³⁴ See G. Thomas Munsterman, Paula L. Hannaford, and G. Marc Whitehead, eds, *Jury Trial Innovations* 151–52 (National Center for State Courts 1997) (discussing preinstructing the jury); *ABA Standards for Criminal Justice Discovery and Trial by Jury* 231 (ABA 3d ed 1996) (mentioning the use of preliminary jury instructions); E. Barrett Prettyman, *Jury Instructions—First or Last?*, 46 *ABA J* 1066 (1960) (suggesting that jury instructions should be given at the start of a trial).

¹³⁵ Lynne ForsterLee, Irwin A. Horwitz, and Martin J. Bourgeois, *Juror Competence in Civil Trials: Effects of Preinstruction and Evidence Technicality*, 78 *J Applied Psych* 14, 18 (1993) (finding that postinstructed jurors were less able to differentiate among the plaintiffs than preinstructed jurors).

¹³⁶ Smith, 76 *J Applied Psych* at 224 (cited in note 128). In the four conditions, participants were given jury instructions before the evidence, after the evidence, both before and after, or never at all. The author found that participants who were instructed both before and after the trial performed significantly better than all others in applying the law to the facts of the case. *Id.*

¹³⁷ Larry Heuer and Steven D. Penrod, *Instructing Jurors: A Field Experiment with Written and Preliminary Instructions*, 13 *L & Human Behav* 409, 424 (1989). Specifically, preinstruction was reported to be more helpful in evaluating the evidence, understanding the law, and applying the law to the facts. *Id.* No differences were found on questions about the clarity of the judge's

A second potential advantage of preinstruction concerns jurors' memory of the evidence. Coherence-based reasoning predicts that preinstruction improves recall by providing jurors with a conceptual framework that contributes to the associative relationship among the facts of the case. The findings on this point have been positive overall, with three out of four experiments demonstrating improved recall.¹³⁸

A third potential advantage of preinstruction pertains to jurors who somehow refrain from imposing preconceived views of the law onto the evidence. Confronted with complicated evidence and unarmed with knowledge of the appropriate conceptual basis for the looming decision, these jurors might experience anxiety and feel ill-prepared for the task. Research in basic psychology shows that performance of difficult tasks is positively related to people's perceptions of their efficacy.¹³⁹ In such situations, people are more likely to base their decisions on various heuristics and stereotypes that could detrimentally affect the case at hand.¹⁴⁰ Preinstruction could thus improve juror performance. Coherence-based reasoning research suggests that

instructions, nor on whether the instructions reduced the confusion about the trial procedure. *Id.* at 425.

¹³⁸ See ForsterLee, Horowitz, and Bourgeois, 78 *J Applied Psych* at 19 (cited in note 135) (finding that preinstruction increases juror recall of probative evidence and decreases recall of nonprobative facts). See also Lynne ForsterLee and Irwin A. Horowitz, *Enhancing Juror Competence in a Complex Trial*, 11 *Applied Cognitive Psych* 305, 314 (1997) (replicating the earlier study with added trial complexity and achieving similar results); Kassin and Wrightsman, 37 *J Personality & Soc Psych* at 1881 (cited in note 128) (finding moderately positive effects). I have not included the work of Elwork and his colleagues in the count of positive findings since they found improvement in only two of the three key factual issues of the trial. See Amiram Elwork, Bruce D. Sales, and James J. Alfini, *Juridic Decisions: In Ignorance of the Law or in Light of It?*, 1 *L & Human Behav* 163, 177 (1977).

Vicki Smith, however, found no improvement in evidence recall due to preinstruction. Smith, 76 *J Applied Psych* at 223 (cited in note 128). Heuer and Penrod also found no improvement in juror evidence recall, but this report should be taken with caution since the study did not include a test for recall. See Heuer and Penrod, 13 *L & Human Behav* at 409 (cited in note 137). The authors' conclusion was based only on jurors' self-reported levels of confidence of recall, a measure that is often infected with metacognitive difficulties and overconfidence. See *id.*

¹³⁹ See, for example, Albert Bandura, *Social Cognitive Theory of Self-Regulation*, 50 *Org Beh & Human Dec Soc Processes* 248, 257-58 (1991) (concluding that confidence in self-efficacy positively affects choices, aspirations, effort, perseverance, and stress levels); Barry J. Zimmerman, *A Social Cognitive View of Self-Regulated Academic Learning*, 81 *J Educ Psych* 329, 331 (1989) (summarizing data suggesting that perceptions of high self-efficacy are positively associated with persistence and achievement in an academic context).

¹⁴⁰ See, for example, Felicia Pratto and John A. Bargh, *Stereotyping Based on Apparently Individuating Information: Trait and Global Components of Sex Stereotypes under Attention Overload*, 27 *J Exper Soc Psych* 26, 44 (1991) (finding that people faced with an unassimilable amount of data tended to rely on sexual stereotypes in judging behavior); Galen V. Bodenhausen and Meryl Lichtenstein, *Social Stereotypes and Information-Processing Strategies: The Impact of Task Complexity*, 52 *J Personality & Soc Psych* 871 (1987) (concluding that when facing a complex judgmental situation, participants were more likely to base their judgments on stereotypes as a way of simplifying the judgment, causing (negative) stereotype-confirming decisions accompanied by high recall of the negative evidence and a neglect of the positive information).

providing jurors with conceptual legal blueprints for their mental models would enhance their sense of competence. The only available evidence to this effect supports the prediction; jurors in the Wisconsin study reported higher levels of perceived competence in tackling their task.¹⁴¹

A potential disadvantage of preinstruction is that the procedure will complicate the trial, especially by making it hard for judges to predict which instructions should be given.¹⁴² Instructions that turn out to be irrelevant in light of the subsequent evidence might somehow affect verdicts. Though theoretically possible, the concern applies only to the infrequent instances where a surprise renders the initial cause of action or criminal charge irrelevant, without ending the trial by way of a directed verdict.¹⁴³ Furthermore, the preliminary instructions will be conveyed as prefatory and explicatory, not as conclusive statements of law. The judge will forewarn jurors that changes are to be expected, and will have ample opportunity to enter revisions.

A related objection is that the opportunity for counsel to submit special requests for preinstruction might complicate preparation and delay trials. But, in most cases, the jurisdiction's pattern jury instructions are not a cause of dispute, and when they are contested, clarification in advance might well be advantageous in helping lawyers better plan their cases. Furthermore, any time spent on determining the content of the preliminary instructions will likely be saved at the post-trial phase. In the Wisconsin field experiment, for instance, the twenty-nine participating judges were specifically asked about these difficulties and denied finding them problematic.¹⁴⁴ Indeed, a majority of these judges asserted their general belief that preinstruction increases the fairness of the trial.¹⁴⁵

Another important concern is that pretrial instructions might cause systematic biases in favor of one side or the other. If, as I believe, erroneous preconceptions are randomly distributed, corrective instructions should result in a decrease in error that is evenly distrib-

¹⁴¹ Heuer and Penrod, 13 *L & Human Behav* at 424 (cited in note 137) (finding that jurors believed preinstruction procedures helped them to be more effective in their responsibilities).

¹⁴² See Vicki L. Smith, *The Feasibility and Utility of Pretrial Instruction in the Substantive Law: A Survey of Judges*, 14 *L & Human Behav* 235, 243-45 (1990). In a survey of the California judges, 71 percent of the judges who expressed their opposition to preliminary instructions cited this concern. Another notable concern was that the practice would waste time and effort. *Id.*

¹⁴³ See FRCP 50 (providing the standard for a directed verdict); FRCrP 29 (same for criminal trials).

¹⁴⁴ Heuer and Penrod, 13 *L & Human Behav* at 426 (cited in note 137).

¹⁴⁵ *Id.* Positive responses were reported also by two judges and several counsel who participated in a small-scale experiment conducted in the Second Circuit. Leonard B. Sand and Steven Alan Reiss, *A Report on Seven Experiments Conducted by District Court Judges in the Second Circuit*, 60 *NYU L Rev* 423, 439-42 (1985) (discussing fourteen cases in which preinstructions were given).

uted. The sole type of instruction for which one might expect an asymmetric effect is the heightened requirement of proof in criminal trials. A preinstruction on the beyond a reasonable doubt standard could be expected to lead jurors to scrutinize the prosecution's evidence more thoroughly, thus lowering the conviction rates. Though mixed, the experimental evidence on this matter generally tends to refute this hypothesis,¹⁴⁶ presumably because most people are relatively familiar with the beyond a reasonable doubt standard, and are thus capable of employing it correctly without the aid of judicial instructions.¹⁴⁷ Moreover, even if preinstruction on the requirements of proof did in fact tip the balance, that would hardly be an undesirable outcome. It would merely mean that criminal defendants would benefit from a fundamental constitutional protection that had previously been misapplied. The fear of systematic changes in the distribution of verdicts due to preinstruction might well be a cause for bureaucratic obstruction, but it hardly makes for a principled objection. It amounts to a perpetuation of erroneous decisions due to misconceived legal rules and jury ignorance of constitutional guarantees.

It has also been suggested that in criminal trials, preliminary instructions might tip the balance in the favor of the prosecution. Since the defense is bound to withhold its strategy until after the presentation of the prosecution's evidence, jurors may initially receive instructions that bear mostly on the prosecution's case. The fear is that this might cause jurors to improperly focus their attention on only half of the picture.¹⁴⁸ Indeed, during the first part of the trial, jurors do fix their attention primarily on the prosecution's case; that is simply because the bulk of the evidence is devoted to that side. For the same reason, jurors concentrate more on the defendant's case during the subsequent part of the trial. There is little reason to believe that ju-

¹⁴⁶ The hypothesis is supported by Kassin and Wrightsman's finding that, of the participants who were given preinstructions on the requirements of proof, 37 percent returned guilty verdicts in a criminal trial, compared to 59 percent among those instructed after the evidence and 63 percent of those who were not instructed at all. Kassin and Wrightsman, 37 *J Personality & Soc Psych* at 1880 (cited in note 128). Vicki Smith, however, found no difference in conviction levels. Smith, 76 *J Applied Psych* at 225 (cited in note 128). In an unpublished experiment, run on a considerably larger group of participants, Steve Read, Chad Snow, and I found no difference in conviction rates.

¹⁴⁷ Research that monitors perceptions of the standard reveals that while assessments vary with the methodology of elicitation, the responses tend to converge in the range of 80 to 90 percent certainty satisfying the standard. See, for example, Reid Hastie, *Algebraic Models of Juror Decision Processes*, in Reid Hastie, ed, *Inside the Juror: The Psychology of Juror Decision Making* 84, 102 (Cambridge 1993) (listing results from numerous empirical studies).

¹⁴⁸ See Wayne R. Lafave, Jerold H. Israel, and Nancy J. King, 5 *Criminal Procedure* § 24.8(a) at 567 (West 2d ed 1999 & Supp 2003) (noting that arguments against substantive preliminary instructions include the possibility that jurors will "improperly focus . . . on only some of the many elements they must consider").

rors' focal point will be affected by the fact that they approach either part of the evidence with a more accurate understanding of the law.

A final concern is that providing a conceptual framework at the beginning of the trial might foster hypothesis-confirming reasoning, which would result in premature decision making.¹⁴⁹ The little available experimental evidence refutes that concern. The findings indicate that preinstructed jurors tend to make their decisions at a later stage than their post-instructed counterparts.¹⁵⁰

The law with respect to preliminary instructions is rather barren. For the most part, judges have broad discretion with respect to the form and timing of jury instructions.¹⁵¹ Preliminary instructions are explicitly included or suggested in procedural law and pattern jury instructions in some jurisdictions,¹⁵² but not in others.¹⁵³ The little case law on the topic has not exhibited any notable objection to the practice.¹⁵⁴ The key to the widespread adoption of substantive preliminary instruction procedure lies mostly in the hands of judges and lawyers.¹⁵⁵

¹⁴⁹ See Heuer and Penrod, 13 L & Human Behav at 414–15 (cited in note 137), citing Reid Hastie, Final Report to the National Institute for Law Enforcement and Criminal Justice (1983). This concern was also cited by a minority of the California judges surveyed by Smith. Smith, 14 L & Human Behav at 244 (cited in note 142).

¹⁵⁰ See Bourgeois, et al, 80 J Applied Psych at 61 (cited in note 128) (finding that preinstructed jurors did not decide the case early); Smith, 76 J Applied Psych at 225 (cited in note 128) (“Contrary to the fears of critics, jurors instructed before trial were significantly more likely to defer their verdict decisions than were those instructed after trial.”).

¹⁵¹ See Kevin F. O'Malley, Jay E. Grenig, and William C. Lee, *Federal Jury Practice and Instructions: Jury Trial* § 7.06 at 504 (West 5th ed 2000 & Supp 2003) (“The federal judge need not give instructions in any particular form . . . so long as the substance of the suggested instruction is fully stated.”).

¹⁵² Both Rule 51 of the Federal Rules of Civil Procedure and Rule 30 of the Federal Rules of Criminal Procedure provide that judges may instruct the jury before or after the arguments, or at both times. Preinstruction would thus be introduced in addition to the statutory requirement. See, for example, Manual of Model Civil Jury Instructions for the District Courts of the Ninth Circuit § 1 (West 2001); Manual of Model Criminal Jury Instructions for the District Courts of the Eighth Circuit § 1 (West 2000); Pennsylvania Suggested Standard Civil Jury Instructions § 1.31 (Pennsylvania Bar Institute 2003); Howard G. Leventhal, *Charges to the Jury and Requests to Charge in a Criminal Case: New York* § 1:03 (Callaghan 2003).

¹⁵³ See, for example, Illinois Pattern Jury Instructions: Civil 5–15 (West 2000); Illinois Pattern Jury Instructions: Criminal § 1.01 (West 4th ed 2000); New York Pattern Jury Instructions—Civil § 1 (Law Co-op 3d ed 1996).

¹⁵⁴ See, for example, *United States v Marsh*, 144 F3d 1229, 1238 (9th Cir 1998) (noting, without criticism, that preliminary jury instructions were given); *United States v Hegwood*, 977 F2d 492, 495 (9th Cir 1992) (holding that faulty preliminary jury instructions were cured by a subsequent correct charge).

¹⁵⁵ A survey administered to 157 Superior Court Judges in California in 1990 revealed that 43 percent of the surveyed judges stated that they favored preinstruction, whereas 57 percent opposed it. Eighty-two percent reported giving procedural instructions before the evidence phase, 74 percent instructed on the requirements of proof, and 34 percent gave preinstructions on substantive law. The survey was sent to 350 of the then-712 Superior Court judges; 45 percent responded. Smith, 14 L & Human Behav at 241–42 (cited in note 142). It should be noted that neither the civil nor the criminal manuals of California Jury Instructions contain pretrial instructions.

As with any bureaucratic change, concerns born of unfamiliarity and fears about administrative feasibility are bound to hinder reform efforts. The best way to allay these anxieties—and, indeed, to verify that they are unwarranted—is by implementing the change in an incremental manner, while carefully monitoring its consequences. Worries of this kind should not thwart improving the fairness and accuracy of trials.¹⁵⁶

B. Atomism, Holism, and the Admissibility of Prejudicial Evidence

The second application of coherence-based reasoning theory to trial procedure concerns a long-standing debate that goes to the heart of evidence law—whether fact-finders evaluate evidence in a holistic or an atomistic manner.¹⁵⁷ This debate pertains to the ever-important and befuddling rule of relevancy and its exceptions due to prejudicial impact.¹⁵⁸ Coherence-based reasoning provides empirical support for the holistic view, but questions the prescriptive corollary advocated by its proponents.

One approach to evidentiary inference is based on mathematical models that capture the probabilistic value of the evidence. Most notable in this line is the reliance on Bayes Theorem, which appears to have attained prominence within mainstream evidence scholarship.¹⁵⁹ Bayesians posit that the probability of an event can be derived algebraically by means of a sequential multiplication of the probabilistic values of the event's constitutive elements.¹⁶⁰ Verdicts are determined

¹⁵⁶ Admittedly, there is no panacea in these matters, and one ought not expect too marked an improvement in jury performance. On the persistence of erroneous beliefs about the law, see Smith, 17 L & Human Behav at 532–35 (cited in note 132).

¹⁵⁷ The debate has been the topic of at least three dedicated conferences. See *Bayesianism and Juridical Proof*, 1 Intl J Evid & Proof 253–360 (1997); *Decision and Inference in Litigation*, 13 Cardozo L Rev 253–1079 (1991); *Symposium: Probability and Inference in the Law of Evidence*, 66 BU L Rev 377–952 (1986), excerpted as Peter Tillers and Eric D. Green, eds, *Probability and Inference in the Law of Evidence: The Uses and Limits of Bayesianism* (Kluwer 1988). The debate also takes up much of the symposium: *New Perspectives on Evidence*, 87 Va L Rev 1491–2081 (2001). The debate can be traced back to Richard Lempert's seminal piece, *Modeling Relevance*, 75 Mich L Rev 1021 (1977) (advocating Bayes Theorem as it relates to evidence processing).

For dispassionate overviews of the debate, see generally Mirjan Damaška, *Atomistic and Holistic Evaluation of Evidence: A Comparative View*, in David S. Clark, ed, *Comparative and Private International Law: Essays in Honor of John Henry Merryman on his Seventieth Birthday* 91 (Duncker & Humblot 1990); William Twining, *Rethinking Evidence: Exploratory Essays* 219–61 (Basil Blackwell 1990).

¹⁵⁸ Federal Rule of Evidence 403 prescribes the exclusion of evidence “if its probative value is substantially outweighed by the danger of unfair prejudice.” The balancing of relevancy and prejudice is central also to the question of admissibility of character evidence, FRE 404(a), and evidence of past crimes, wrongs, or acts, FRE 404(b).

¹⁵⁹ See, for example, Richard O. Lempert, Samuel R. Gross, and James S. Liebman, *A Modern Approach to Evidence: Text, Problems, Transcripts and Cases* 228–39 (West 3d ed 2000) (applying Bayesian analysis to explain basic concepts in evidence).

¹⁶⁰ See, for example, David H. Kaye, *Introduction: What Is Bayesianism?*, in Tillers and

by comparing the posterior likelihood produced by the Bayesian computation with the numerical value assigned to the respective standard of proof.¹⁶¹ Mathematical approaches are also described as *atomistic*, in that each piece of evidence is individually evaluated, quantified, and entered at face value into the calculus of inference.¹⁶²

Bayes Theorem relies heavily on syntactic assumptions, the most important of which is that each piece of evidence is evaluated on its own terms, and is affected neither by the other pieces of evidence (unless the items are substantively interdependent) nor by the conclusion of the process. This assumption relies in turn on a conception of uni-directionality in human reasoning: inferences flow from the individual pieces of evidence toward a computed judgment, but the perception of the evidence is in no way affected in the reverse direction. In other words, the evaluation of the evidence precedes, and is thus entirely exogenous to, the process of making the judgment.

The alternative position conceives of juridical inference making in *holistic* terms. According to this view, evidence is evaluated not as isolated pieces, but rather in large cognitive structures, most familiarly in the form of narratives, stories, or global accounts. The sufficiency of holistic proof is evaluated not on the probabilistic value of a single focal hypothesis, but rather, as advocated by Ronald Allen, by a global comparative assessment of the competing evidential accounts.¹⁶³ This “Bayesian-skeptical”¹⁶⁴ approach proposes that legal fact-finding hinges on the *relative plausibility* of the vying explanations offered at trial. In civil cases, the fact-finder determines the facts by simply identifying

Green, eds, *Probability and Inference in the Law of Evidence* 1 (cited in note 157).

¹⁶¹ See, for example, Lempert, Gross, and Liebman, *A Modern Approach to Evidence* at 237 (cited in note 159); Reid Hastie, *Introduction*, in Hastie, ed, *Inside the Juror* 3, 15 (cited in note 147).

¹⁶² For similar reasons, mathematical models have also been called meter models. As Lola Lopes explains, “The meter model is based on the conception that inside the head are one or more mechanisms that continuously track, and can read out, one or more values that reflect the juror’s current evaluation of the evidence.” Lola Lopes, *Two Conceptions of the Juror*, in Hastie, ed, *Inside the Juror* 255, 255 (cited in note 147).

¹⁶³ See Ronald J. Allen, *Rationality, Algorithms, and Juridical Proof: A Preliminary Inquiry*, 1 *Intl J Evid & Proof* 254 (1997); Ronald J. Allen, *Factual Ambiguity and a Theory of Evidence*, 88 *Nw U L Rev* 604 (1994) (addressing three ways in which the relative plausibility theory is superior to Bayesian theory). For a recent formulation, see Ronald J. Allen and Brian Leiter, *Naturalized Epistemology and the Law of Evidence*, 87 *Va L Rev* 1491, 1527–37 (2001) (explaining the emergence of relative plausibility theory as a response to Bayesian theory). For Bayesian-based critiques, see Dale A. Nance, *Naturalized Epistemology and the Critique of Evidence Theory*, 87 *Va L Rev* 1551, 1595–1616 (2001) (claiming Allen and Leiter did not sufficiently realize the role for Bayesianism in the naturalized epistemology discussion); Richard D. Friedman, “*E*” is for *Eclectic: Multiple Perspectives on Evidence*, 87 *Va L Rev* 2029, 2040–48 (2001) (elucidating Bayesian responses to Allen and Leiter).

¹⁶⁴ For use of the term Bayesian-skeptical, see Richard D. Friedman, *Economic Analysis of Evidentiary Law: An Underused Tool, an Underplowed Field*, 19 *Cardozo L Rev* 1531, 1535 (1998).

the most plausible account. In criminal trials, she will convict only if the prosecution provides an account of guilt that leaves no room for a plausible account of innocence. A central claim of relative plausibility theory is that it best comports to what jurors actually do.¹⁶⁵ The empirical work most often cited in support for holism is Nancy Pennington and Reid Hastie's "story model" research.¹⁶⁶

According to the story model, people make sense of complicated bodies of evidence by constructing narratives, rather than by aggregating measures of the strength of the raw evidence. Stories are structured around episodes held together by certain types of causal and physical relationships that explain the actors' conduct and mental states. A pervasive type of narrative structure is based on schemas of human action: initiating events cause characters to respond in a variety of ways and to form goals that then motivate subsequent actions. The acceptance of a story is determined by its global strength based on four certainty principles—coverage, coherence, uniqueness, and goodness-of-fit. Of the stories constructed from the evidence presented at trial, a juror adopts as the best explanation for the events the one that best meets these certainty principles.¹⁶⁷

An interesting feature of the evidence debate is that most participants tend to proclaim superiority on both descriptive and prescriptive grounds.¹⁶⁸ The shared assumption seems to be that an atom-

¹⁶⁵ See Allen, 88 Nw U L Rev at 604 (cited in note 163) ("[C]ognitive psychology generally, and research into jury decisionmaking specifically, [] have made it rather plain that virtually no one thinks as the [Bayesian] legal theory requires."); Allen and Leiter, 87 Va L Rev at 1528 (cited in note 163).

¹⁶⁶ See generally Reid Hastie and Nancy Pennington, *Explanation-Based Decision Making*, in Terry Connolly, Hal R. Arkes, and Kenneth R. Hammond, eds., *Judgment and Decision Making: An Interdisciplinary Reader* 212 (Cambridge 2d ed 2000); Nancy Pennington and Reid Hastie, *The Story Model for Juror Decisionmaking*, in Hastie, *Inside the Juror* 192 (cited in note 147).

¹⁶⁷ For experimental results, see Nancy Pennington and Reid Hastie, *Explaining the Evidence: Tests of the Story Model for Juror Decision Making*, 62 J Personality & Soc Psych 189 (1992) (finding that story model constructions were more likely to be used when jurors are asked to make a global judgment at the end of the case, but linear models were more likely to be used when jurors are asked to make judgments after each evidence block); Nancy Pennington and Reid Hastie, *Explanation-Based Decision Making: Effects of Memory Structure on Judgment*, 14 J Exper Psych: Learning, Memory & Cognition 521 (1988) (finding that manipulating the order in which evidence was presented shifted verdicts toward the most easily constructed story); Nancy Pennington and Reid Hastie, *Evidence Evaluation in Complex Decision Making*, 51 J Personality & Soc Psych 242 (1986) (finding that jurors organize evidence into causal relationships when asked to decide on a verdict). The story model was preceded by work by Lance Bennett and Martha Feldman. See W. Lance Bennett and Martha S. Feldman, *Reconstructing Reality in the Courtroom: Justice and Judgment in American Culture* 41–90 (Rutgers 1981).

¹⁶⁸ Compare Allen and Leiter, 87 Va L Rev at 1507–10 (cited in note 163) (critiquing Bayes Theorem on both conceptual and empirical grounds and charging that computational complexity undermines its plausibility), with Nance, 87 Va L Rev at 1599–1602 (cited in note 163) (defending Bayes Theorem from the computational complexity argument on both descriptive and prescriptive grounds).

As an exception to this tendency, some Bayesian-oriented evidence scholars have been

istic conception warrants an interventionist set of rules tailored to regulate the admissibility of each and every piece of evidence, whereas a holistic conception justifies only minimal interference in the flow of evidence—an evidentiary regime known as “free proof.”¹⁶⁹

Coherence-based reasoning research speaks to both the descriptive and prescriptive dimensions of the evidence debate. The experimental results contest the descriptive validity of the atomistic approaches by showing that participants defy the syntactic rules of unidirectional inference. The perception of evidence changes as the pieces interact with the emerging verdict and indirectly also with one another; human reasoning progresses bidirectionally, from evidence to conclusion and back to evidence.¹⁷⁰

These experimental results favor the holistic approach of relative plausibility theory. The fact that the evidentiary items on each of the respective sides are found to shift together suggests that processing is done in a global, rather than an atomistic, manner. The notion of relative plausibility is also distinctly present in the skewed nature of the mental models, in which the evidence for one side is perceived to be highly plausible while the other subset is deemed weak. Moreover, the confidence levels reported in the experimentation were strongly related to the difference between the ratings of the inculcating and exculpating evidence. In other words, the higher the discrepancy between the plausibility of the verdicts, the higher the levels of confidence reported.¹⁷¹ This ordinal comparison between the accepted and the rejected version of facts is unique to relative plausibility theory; Bayesian theories look only to a cardinal comparison between the facts of the single focal hypothesis and the standard of proof.¹⁷²

rather conciliatory on the question of empirical validity. See Richard Lempert, *Of Flutes, Oboes, and the As If World of Evidence Law*, 1 Intl J Evid & Proof 316, 320 (1997) (emphasizing that applying Bayes Theorem to evidence law does not depend on its descriptive accuracy); Richard D. Friedman, *Answering the Bayesioskeptical Challenge*, 1 Intl J Evid & Proof 276, 291 (1997) (same); Lempert, 75 Mich L Rev at 1023, 1056 (cited in note 157) (same).

¹⁶⁹ See Allen, 88 Nw U L Rev at 632 n 88 (cited in note 163) (stating that free proof is implicit in his proposed theory of evidence); Allen and Leiter, 87 Va L Rev at 1536 (cited in note 163) (arguing that liberal standards of admission support the relative plausibility theory). The argument is also made explicitly by one of Allen’s students. See Michael S. Pardo, Comment, *Juridical Proof, Evidence, and Pragmatic Meaning: Toward Evidentiary Holism*, 95 Nw U L Rev 399, 441 (2000) (advocating a shift to a free proof system based on the holistic account).

¹⁷⁰ While this experimental finding belies the syntactic rules of uni-directional inference, it does not necessarily contradict the Bayesian tenet of calculating posterior odds by updating the likelihood ratios. It is theoretically possible that evidence undergoes coherence shifts and then gets calculated by means of a Bayesian updating. Though I believe this possibility to be rather remote, our experimental results do not prove or disprove any specific algorithm.

¹⁷¹ See note 60 and accompanying text.

¹⁷² It should also be noted that relative plausibility theory is claimed to be based on connectionist conceptions of cognition. See Allen and Leiter, 87 Va L Rev at 1528 (cited in note 163), referencing Paul Thagard, *Conceptual Revolutions* 25–27 (Princeton 1992).

The experimental results explain some empirical phenomena that Bayesian researchers observe but cannot explain. In a frequently cited experimental investigation by Bayesian advocates, David Schum and Anne Martin found it “[q]uite startling” that unpredicted interdependence appeared between evaluations of seemingly unrelated pieces of evidence.¹⁷³ Similarly, in another experiment testing for Bayesian updating, experimenters were surprised to find that ratings of prosecution evidence increased after the presentation of the defendant’s weak alibi.¹⁷⁴ These findings are readily explainable as coherence effects that result from constraint satisfaction processing.¹⁷⁵

Coherence research fits with the story model. Both approaches posit that evidentiary conclusions are not derived from mathematical computations of the independent values of raw evidence. Inferences, rather, are based on constructed representations of coherence, and it is these constructed representations that ultimately determine the verdicts. Coherence research overcomes an important limitation of the story model. As indicated by its name, the story model proposes that the representation of evidence bears a distinct narrative structure. Pennington and Hastie’s results showed that participants’ stories were centered on a narrative that captured the intentional and causal scheme of the defendant’s behavior—intentions, psychological states,

¹⁷³ David A. Schum and Anne W. Martin, *Formal and Empirical Research on Cascaded Inference in Jurisprudence*, 17 L & Soc’y Rev 105, 144 (1982). Schum and Martin found that participants interpreted contradictory testimony as either probatively valueless or even corroborative, and “double counted” corroboratively redundant testimony. Consistent with coherence-based reasoning, these findings were observed only when participants were instructed to make overall evaluations of the evidence; they were not obtained when the evidence was assessed piecemeal. See *id.*

¹⁷⁴ See Brian C. Smith, et al, *Jurors’ Use of Probabilistic Evidence*, 20 L & Human Behav 49 (1996).

¹⁷⁵ It should be noted, however, that the findings of coherence-based reasoning are not radically discordant with some of the more nuanced and sophisticated views offered by Bayesian scholars. For example, Richard Lempert has approvingly discussed the contribution of the story model to evidence law. Richard Lempert, *Telling Tales in Court: Trial Procedure and the Story Model*, 13 Cardozo L Rev 559 (1991) (exploring the utility of the story model in each aspect of the trial process). Similarly, Richard Friedman has espoused a hybrid “storytelling Bayesian approach,” which is not very dissimilar to coherence-based reasoning. Friedman explains that

an observer might at any time make a set of probability assessments that are inconsistent. If the inconsistency is a glaring one, the cognitive dissonance will be apparent, and the observer will adjust the assessments to bring them more closely into line. This adjustment might involve altering any probability assessment, including a prior probability. The tendency will be to move towards equilibrium, in which there are no inconsistencies, but new information might disrupt the situation and start the process again.

Friedman, 1 Intl J Evid & Proof at 289 (cited in note 168) (footnote omitted). Both Lempert and Friedman have also relaxed the Bayesian precept on separately updating the likelihood ratios of every piece of evidence, suggesting instead that evidence can be “chunked,” or “batched” in ways that closely resemble coherence-based processing. See *id.* at 288; Lempert, 1 Intl J Evid & Proof at 317 (cited in note 168).

goals, and motivations.¹⁷⁶ But, in a range of evidentiary situations, the material facts “may concern a situation or state of affairs rather than a sequence of events.”¹⁷⁷ A narrative of human intentionality is of little relevance to negligence cases where the contested issue is a failure in appreciating a risk, to identification cases, or to cases in which the material facts concern the physical conduct of the defendant, the quality of a product, the extent of damages, and the like.¹⁷⁸

With this concern in mind, the *Jason Wells* case was designed to contain evidence that is entirely physical, circumstantial, and disjointed. Our findings indicate that restructuring of cognitive representations is not limited to stories that hinge on human intentionality. Coherence-driven representations are found in the absence of such narrative possibilities.¹⁷⁹ Furthermore, by relating coherence effects to constraint satisfaction theory, coherence-based reasoning provides a deeper theoretical explanation of holistic processing, thus enabling its extension to other realms of reasoning.¹⁸⁰ While narrative structures might not be essential for holistic processing, coherence shifts are likely to be particularly pronounced in their presence.¹⁸¹

The evidence debate has recently been joined by the Supreme Court in the case of *Old Chief v United States*.¹⁸² Appellant Old Chief had previously been sentenced to five years’ imprisonment for committing an assault that caused serious bodily injury. The current case concerned an altercation in which he allegedly fired a gun. The

¹⁷⁶ Pennington and Hastie, *The Story Model for Juror Decision Making* at 196 (cited in note 166) (“Stories involve human action sequences connected by relationships of physical causality and intentional causality between events.”).

¹⁷⁷ Twining, *Rethinking Evidence* at 225 (cited in note 157).

¹⁷⁸ The story model’s applicability is equally questionable in tasks of evidence integration outside the realm of legal decisions, such as judgments about historical and physical events, as well as scientific knowledge. For applications of constraint satisfaction-based processing to such domains, see Keith J. Holyoak and Paul Thagard, *Mental Leaps: Analogy in Creative Thought* (MIT 1995); Thagard, *Coherence in Thought and Action* (cited in note 25).

¹⁷⁹ The extension of holistic processing to identity cases is noteworthy, in that they account for a great deal of convictions of innocent people. For errors in adjudication of identity cases, see generally Barry Scheck, Peter Neufeld, and Jim Dwyer, *Actual Innocence: Five Days to Execution and Other Dispatches from the Wrongly Convicted* (Doubleday 2000).

¹⁸⁰ It should also be noted that while the story model is compelling on both theoretical and methodological grounds, it did not include a baseline comparison. By enabling a comparison between participants’ representations of the evidence at two points in time, coherence-based reasoning research provides more direct proof of the reconstruction of evidence.

¹⁸¹ See Melanie C. Green and Timothy C. Brock, *The Role of Transportation in the Persuasiveness of Public Narratives*, 79 *J Personality & Soc Psych* 701 (2000) (showing that evaluations of facts and beliefs are related to the story’s narrative transportability); James F. Voss, Jennifer Wiley, and Rebecca Sandak, *On the Use of Narrative as Argument*, in Susan R. Goldman, Arthur C. Graesser, and Paul van den Broek, eds, *Narrative Comprehension, Causality, and Coherence: Essays in Honor of Tom Trabasso* 235 (Lawrence Erlbaum 1999) (demonstrating the effect of narrative structure on simulated juror verdicts).

¹⁸² 519 US 172 (1997).

charges included a violation of a law forbidding convicted felons from possessing firearms.¹⁸³ Old Chief feared that if the jury were to learn of the full extent of his previous assault, it would be more prone to convict him on the current charges. He thus moved for an order prohibiting the prosecution from revealing the nature of the previous conviction and the sentence imposed on him; instead, he offered to stipulate to a prior conviction that would meet the standard of the “felon in possession” law.¹⁸⁴ The trial court heeded the prosecution’s insistence on describing the nature of the previous conviction to the jury in full. Old Chief was convicted, and his appeal was rejected by the Ninth Circuit Court of Appeals. Writing for a five-justice majority, Justice Souter reversed the conviction, finding that in this specific case, the stipulation did not compromise the prosecution’s case.¹⁸⁵

Despite the narrow holding, this case has drawn a great deal of attention due to the potential effects of its broad dictum on evidence doctrine. First, Justice Souter openly rejected the atomistic view. An item of evidence should not be considered an “island,” whose “own probative value and unfairly prejudicial risk [are] the sole reference points” of admissibility.¹⁸⁶ The opinion then goes on to endorse a distinct version of holism:

Unlike an abstract premise, whose force depends on going precisely to a particular step in a course of reasoning, a piece of evidence may address any number of separate elements, striking hard just because it shows so much at once; the account of a shooting that establishes capacity and causation may tell just as much about the triggerman’s motive and intent. Evidence thus has force beyond any linear scheme of reasoning, and as its pieces come together a narrative gains momentum, with power not only to support conclusions but to sustain the willingness of jurors to draw the inferences, whatever they may be, necessary to reach an honest verdict.¹⁸⁷

This account is consistent with coherence-based reasoning research and other holistic approaches. Each piece of evidence is seen as having more implications than its own isolated probativeness in a linear chain of inference; the pieces come together to form a narrative that is stronger than the sum of its parts. The opinion should be cele-

¹⁸³ See 18 USC § 922(g) (2000). For factual findings, see *Old Chief*, 519 US at 174–77.

¹⁸⁴ See *Old Chief*, 519 US at 175. Appellant requested the trial court to “instruct the jury that he has been convicted of a crime punishable by imprisonment exceeding one (1) year.” *Id.*

¹⁸⁵ Justice Souter’s opinion was joined by Justices Stevens, Ginsburg, Breyer, and Kennedy. *Id.* at 173.

¹⁸⁶ *Id.* at 182.

¹⁸⁷ *Id.* at 187.

brated for taking a step toward a realistic reckoning with the capabilities and limitations of human cognition.¹⁸⁸

But the Court did more than embrace a particular concept of human cognition. It followed its endorsement with a matching prescription:

[T]he prosecution may fairly seek to place its evidence before the jurors, as much to tell a story of guiltiness as to support an inference of guilt, to convince the jurors that a guilty verdict would be morally reasonable as much as to point to the discrete elements of a defendant's legal fault.¹⁸⁹

The doctrinal implication of the opinion lies in its innovative notion of “narrative relevance.”¹⁹⁰ The Court suggested that relevance—and thus, presumably, also admissibility—should not depend solely on conventional tests of idiosyncratic probativeness.¹⁹¹ Relevance, according to the Court, can also be based on how the piece of evidence interacts with the other evidence and how it contributes to the argument's overall narrative force. In the ever-elusive domain of balancing probativeness against prejudicial impact, the Court seems to be shifting the ground toward a more lax inclusion of potentially prejudicial evidence, under which prosecutors may “tell a story of guiltiness.”

To justify this approach, Justice Souter explained that since “[j]ury duty is usually unsought and sometimes resisted,” telling a colorful story helps to secure jurors' “obligation to sit in judgment.”¹⁹²

¹⁸⁸ There is a general sense among legal psychologists that the legal system harbors unrealistic conceptions of human cognition. There is considerable doubt as to jurors' ability to accurately gauge their own susceptibility to bias, to follow judicial instructions to disregard inadmissible evidence, to consider admissible evidence only for a designated purpose, to accurately detect liars, and to be capable of determining truth from scientific evidence. See, for example, Roselle L. Wissler and Michael J. Saks, *On the Inefficacy of Limiting Instructions: When Jurors Use Prior Conviction Evidence to Decide on Guilt*, 9 L & Human Behav 37 (1985); Richard O. Lempert, *Built on Lies: Preliminary Reflections on Evidence Law as an Autopoietic System*, 49 Hastings L J 343 (1998) (“My thesis is that our system of evidentiary rules . . . works in part because it often falsely portrays reality.”). Indeed, Lempert has celebrated *Old Chief* on these grounds. See Richard O. Lempert, *Narrative Relevance, Imagined Juries, and a Supreme Court Inspired Agenda for Jury Research*, 21 SLU Pub L Rev 15 (2002) (arguing that the *Old Chief* Court complicates the questions that must be answered in determining the role of human frailties in establishing evidence admissibility).

The Court's decision is an exception to the other applications discussed in this Article since it is not borne by rationalist assumptions. On the contrary, the Court's endorsement of holism breaks from a priori conceptions about juror rationality. The problem, as I discuss below, lies with the normative implications of the issue.

¹⁸⁹ *Old Chief*, 519 US at 188.

¹⁹⁰ See Lempert, Gross, and Liebman, *A Modern Approach to Evidence* at 251–54 (cited in note 159) (explaining the relationship between narrative relevance and demonstrative evidence).

¹⁹¹ For the conventional test, see FRE 401 (“‘Relevant evidence’ means evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.”).

¹⁹² *Old Chief*, 519 US at 187–88.

Next, he argued that presenting the evidence in all its particularity is needed to satisfy the jurors' expectations about what proper proof should be.¹⁹³ Both of these reasons are wanting and beside the point.¹⁹⁴ Rather, the opinion is based primarily on the same enmeshment of the descriptive with the prescriptive that pervades the academic debate.¹⁹⁵ Underlying the opinion is the seemingly straightforward rationale that the prosecution should be allowed to present its case in a manner that corresponds to the holistic way in which jurors process evidence.

As the coherence experiments show, sufficiently strong pieces of evidence can affect the entire mental model of the case through indirect influences on other variables. For example, describing a defendant as either benevolent or malevolent changed participants' views of a range of unrelated variables, including whether the Internet is analogous to a newspaper. Similarly, the introduction of DNA evidence in the *Jason Wells* theft case affected judgments of numerous other pieces of evidence, including the veracity of an eyewitness's testimony and beliefs about the accuracy of eyewitness identifications in general.¹⁹⁶ The holistic outcome is, after all, an integration of the atoms of evidence that enter the mix.

Governing the admission of evidence by a regime of free proof because it seems to comport with holistic processing could undermine the regulatory purpose of evidence law. This superficial equation of *is*

¹⁹³ See *id.* at 188.

¹⁹⁴ As for the first argument, the Court seems to adopt a highly distrustful view of the jury. If, as the Court suggests, a juror might lose interest in criminal cases, which tend to be inherently interesting, because some especially interesting fact is left out, then one ought to reconsider the trust we place in the institution of the jury. What can be said about jury duty on cases that lack drama altogether? We nonetheless take seriously jury verdicts in lifeless cases involving violations of tax laws, building codes, and other dreary technical issues. We seem to believe that jurors are capable of sitting through such cases and paying due attention without any particular form of stimulation.

The argument that jurors ought to be presented with evidence merely to fulfill their expectations is similarly problematic. Satisfying jurors' expectations by providing them with evidence that would otherwise be ruled prejudicial seems to be a perversion of the normative role of evidence law. Furthermore, the role of the juror is highly regulated and restricted in ways that are bound to frustrate curious minds. Jurors are regularly deprived of large and crucial segments of evidence: they are forbidden to ask the witnesses questions and to explore the scene of the crime, they are often deprived of the testimony of the criminal defendant, they are not allowed to take into account anticipated sentences, and much more. One ought to wonder whether satisfaction of jurors' expectations warrants admitting potentially prejudicial evidence.

For critical analyses of the Court's reasoning, see D. Michael Risinger, *John Henry Wigmore, Johnny Lynn Old Chief, and "Legitimate Moral Force": Keeping the Courtroom Safe for Heartstrings and Gore*, 49 *Hastings L J* 403, 455–58 (1998) (characterizing the opinion as falling into the "near mystical excess of some contemporary narrative theorists"); James Joseph Duane, *"Screw Your Courage to the Sticking-Place": The Roles of Evidence, Stipulations, and Jury Instructions in Criminal Verdicts*, 49 *Hastings L J* 463 (1998) (arguing that, in *Old Chief*, the Court's supposed justifications of narrative integrity and evidentiary richness are implausible).

¹⁹⁵ See text accompanying notes 168–69.

¹⁹⁶ See Part I.B.5.

with *ought* bypasses the basic principle of seeking truth and balancing it with other policy considerations. Thus, for example, prosecutors are normally barred from presenting evidence of a defendant's criminal past not because the jury would find the evidence useless for the construction of a "story of guiltiness."¹⁹⁷ We do so precisely because juries tend to find these kinds of evidence all too useful, leading to verdicts that are likely to be driven by questionable judgments of the defendant's character rather than a fair-minded assessment of the conduct in question.

Research in both attribution theory and stereotyping strongly supports this proposition, showing that people harbor chronically activated blueprints for the judgment of other people. Past conduct, perceived character traits, and group identity are centrally embedded in these mental models, and thus lead readily and often uncritically to corresponding judgments.¹⁹⁸ Coherence-based reasoning suggests that these judgments will be accompanied by coherence shifts in the perception of the evidence. The appropriate test for admissibility of potentially prejudicial evidence should continue to be the one endorsed by the Advisory Committee of Federal Rules of Evidence: whether the specific piece of evidence has an undue tendency to suggest a decision on an improper basis.¹⁹⁹

The Court's proposition warrants attention because it extends beyond questions of the admissibility of predicate offenses in the face of a defendant's stipulation. The language of the opinion could open the door to broader admissibility of other types of evidence, including emotionally arousing information. Under Justice Souter's proposed regime, graphic depictions of "heartstrings and gore" could be admitted more liberally because such evidence adds richness of detail to the story of guiltiness.²⁰⁰

We should be wary of this doctrinal development. The coherence experiments showed that affective dimensions can have a strong impact on putatively rational decision making. Affect-laden outcomes sway mental models so that the evidence is perceived to cohere with

¹⁹⁷ See text accompanying note 189.

¹⁹⁸ For examples of such research, see Daniel T. Gilbert, *Ordinary Personology*, in Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, eds, 2 *The Handbook of Social Psychology* 89 (McGraw-Hill 4th ed 1998); Susan T. Fiske, *Stereotyping, Prejudice and Discrimination*, in Gilbert, Fiske, and Lindzey, eds, 2 *The Handbook of Social Psychology* 357. For an application of constraint satisfaction models to stereotyping, see Kunda and Thagard, 103 *Psych Rev* 284 (cited in note 35).

¹⁹⁹ See FRE 403, Advisory Committee's Note.

²⁰⁰ Such evidence also meets Justice Souter's other objectives of sustaining jurors' willingness to determine the case and convincing them of the moral reasonableness of a conviction. The term "heartstrings and gore" is borrowed from Michael Risinger, 49 *Hastings L J* at 403 (cited in note 194).

the verdict, while maintaining the subjective belief that the decision was properly based on the evidence.²⁰¹ The Federal Rules of Evidence provide sound guidance by drawing attention to the fact that prejudicial evidence is commonly of an emotional nature. Indeed, the effects of emotional influences on judgments of blame are clearly demonstrated in simulations of criminal trials. Studies have shown that emotionally arousing evidence can lead to higher levels of conviction, even when the evidence has no probative value for the issue at hand.²⁰² As a result, litigants' fates are adversely affected by the emotional baggage that comes along with nonprobative, or weakly probative, evidence.

Precisely because the holistic account of human cognition is correct, its prescription is wrong. Due to the connectionist nature of the decision-making process, extra vigilance is needed to contain the adverse effect of potentially prejudicial evidence—a task best performed by an atomistic analysis. Litigants ought to be allowed to tell their stories in court, but preserving the narrative integrity of the stories should not come at the expense of eroding the safeguards against the infiltration of prejudicial evidence.

C. Debiasing Criminal Jurors

The next application concerns the way in which fact-finders reach verdicts in criminal cases. The findings of coherence shifts suggest that even mild perceptions of evidence will polarize and thus seem strongly probative. In this way, coherence effects can amplify the probativeness of ambiguous evidence, elevating the confidence levels of some jurors to beyond a reasonable doubt. Such an effect undermines the protective function of the heightened standard of proof. I suggest

²⁰¹ Participants betting on a horse displayed overconfidence in assessing the likelihood of winning, and participants playing the role of an attorney displayed optimistically distorted perceptions of the case. See text accompanying notes 91–93.

²⁰² For example, Douglas, Lyon, and Ogloff have demonstrated that showing participants photographs of a murder victim's body results in higher levels of conviction than when the state of the body was described verbally in a medical examiner's report, despite the irrelevance of the state of the body to the question at trial (the defense was based on alibi testimony). The heightened level of convictions was found to be mediated by the participants' reported emotional states. See Kevin S. Douglas, David R. Lyon, and James R. Ogloff, *The Impact of Graphic Photographic Evidence on Mock Jurors' Decisions in a Murder Trial: Probative or Prejudicial?*, 21 *L & Human Behav* 485 (1997). Other research supports this finding. See, for example, James R.P. Ogloff and Neil Vidmar, *The Impact of Pretrial Publicity on Jurors: A Study to Compare the Relative Effects of Television and Print Media in a Child Sex Abuse Case*, 18 *L & Human Behav* 507 (1994) (showing that jurors were more likely to believe victim testimony if they had previously heard about the story in the media); Denise H. Whalen and Fletcher A. Blanchard, *Effects of Photographic Evidence on Mock Juror Judgment*, 12 *J Applied Soc Psych* 30, 33–38 (1982) (demonstrating the effect of emotions on the determination of damages in a tort case). For a more sanguine approach to the effect of emotionality on juror decisions, see Eric A. Posner, *Law and the Emotions*, 89 *Georgetown L J* 1977, 1999–2000 (2001) (arguing that emotions such as disgust and anger will not necessarily interfere with jurors' ability to deliberate effectively).

correcting for this problem by introducing a jury instruction to moderate the effects of coherence shifts.

Evidence has no meaningful independent existence but for the decision-maker's perception of it. Thus, to understand how verdicts are decided, one needs to understand how evidence is processed and represented in the decision-maker's mind. Recall that at the vignette phase of the *Jason Wells* experiments, the evidence was indistinct and nonprobative of the defendant's guilt.²⁰³ By the point of decision, however, the perceptions of the evidence had shifted toward skewed mental models that were strongly probative, and the decisions were backed by high levels of confidence.

In civil trials, where the standard of proof is located near the point of indifference, the impact of coherence shifts on determinations of liability should be quite limited. An assessment that falls slightly under or over the mid-point should suffice no less than an extreme assessment in determining the corresponding verdict. Coherence shifts polarize the reasons for deciding for either side, but they should not make a substantial impact on the distribution of verdicts.

In criminal trials, however, the polarization can affect the distribution of verdicts. There is no indication that the participants who voted to convict Jason Wells violated the heightened standard of proof. The evidence, as they perceived it, was strong and unambiguous; they strongly agreed with the inculpatory evidence and strongly disagreed with the exculpatory evidence, reporting high confidence in their judgments.²⁰⁴ Yet, virtually the same evidence was deemed weak and nonprobative before the coherence shifts occurred. These shifts are responsible for the amplification of the weak evidence into evidence demonstrating guilt beyond a reasonable doubt.²⁰⁵ This conclusion is also supported by the finding that the participants whose evidence ratings shifted most intensely were also most confident about their decisions.²⁰⁶ For some of the participants who decided to convict, this bolstered confidence was apparently enough to surpass the be-

²⁰³ See Figure 4 and text accompanying note 58.

²⁰⁴ See text accompanying note 60.

²⁰⁵ The reported ratings represent the mean value for each of the groups of participants, and thus they conceal the true distribution of ratings. One can easily imagine that the coherence shifts made no difference for participants who initially agreed strongly with the inculpatory evidence, nor did they affect those who firmly agreed with the exculpatory evidence. These participants would vote to convict and acquit respectively, regardless of the coherence shifts. Coherence shifts, however, are bound to make a big difference for the presumably sizeable group of people who started off around the point of indifference or with moderate inclinations toward conviction and ultimately decided to convict. Absent the cognitive transformation, these participants would (and should) have decided, perhaps reluctantly, to acquit. They convicted nonetheless due to the extra measure of polarization that drove their mental models and confidence levels above the standard of proof threshold.

²⁰⁶ See text accompanying note 61.

yond a reasonable doubt level. In other words, coherence shifts have the effect of undercutting the protection promised by the heightened standard of proof.²⁰⁷

Assuming that coherence shifts do in fact contribute to the rate of convictions, the normative implications of this phenomenon are not obvious. The participants' decisions to convict were, after all, made in the normal manner in which people make decisions. There was nothing illicit or prejudiced about these verdicts. The real dilemma is whether or not to introduce a moderating mechanism into jury deliberation.

The idea of introducing a moderating intervention is bound to meet resistance. It rubs against the intuition that the legal system should refrain from disrupting jurors' decision-making processes, especially in the absence of any untoward influence. One of the justifications for the use of juries is the premium we place on their common sense and reliance on everyday reasoning.²⁰⁸ Juries, then, should be left to determine verdicts by their normal decision-making processes. While attractive on its face, this deferential approach is unjustifiable. As described above, verdicts are loosely based on the initial perceptions of the evidence, but these perceptions are impacted in turn by the emerging verdicts. This transformation enables choice, but also entails a distortion of the perception of the evidence. The strong probativeness of the mental model is an artifact of the cognitive process, rather than a property of the evidence itself. It is questionable whether we owe allegiance to this inflated sense of coherence.

Furthermore, the phrasing of the proposed instruction—"please take some time to seriously consider the possibility that the opposite side has a better case"²⁰⁹—hardly amounts to an unwarranted disruption or contamination of the decision-making process. The need for the intervention stems from the fact that once people adopt a skewed

²⁰⁷ An increase in the likelihood that a single juror will decide to convict is bound to carry over to decisions by juries. It has been shown repeatedly in both experimental and naturalistic settings that jury verdicts are highly contingent on the predeliberation distribution of votes. In close cases, a shift of as few as one or two jurors can bear a strong impact on the distribution of jury verdicts. See Harry Kalven, Jr. and Hans Zeisel, *The American Jury* (Little, Brown 1966) (arguing, in part, that the majority opinion predeliberation will be equal to the final verdict); Marla Sandys and Ronald C. Dillehay, *First-Ballot Votes, Predeliberation Dispositions, and Final Verdicts in Jury Trials*, 19 *L & Human Behav* 175 (1995) (confirming, mostly, the Kalven and Zeisel findings). On experimentally based findings, see, for example, Dennis J. Devine, et al., *Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups*, 7 *Psych, Pub Pol & L* 622 (2001) (reviewing 206 major studies conducted between 1955 and 1999 involving jury decision making).

²⁰⁸ See Mirjan R. Damaška, *Evidence Law Adrift* 27 (Yale 1997) (arguing that juries follow a decision process with which they are familiar, and which is similar to one used in "ordinary life and personal affairs").

²⁰⁹ See text accompanying note 105.

mental model, they do not spontaneously consider the opposite alternatives. The moderating intervention prods jurors to engage in a mental step that is quite consistent with the expectation that the jury engage in a diligent and thorough consideration of all aspects of the case.

Prosecutors might argue that it is not undesirable that coherence effects lead to heightened levels of conviction. Given the common law's tendency to adapt and gravitate toward optimal social equilibria,²¹⁰ one could infer that the extant conviction levels are the product of such an adaptation. The thrust of the objection, then, is that if we debias jurors, we ought to readjust the system to make it more amenable to convictions; the most natural way to do so is by lowering the standard of proof.²¹¹ This objection makes little sense; it is doubtful that any aspect of the criminal justice system can be traced historically to, or normatively justified by, the purpose of maintaining a predetermined rate of conviction. The standard of proof has a distinct lineage and function of providing a calculus for trading off the costs of erroneous judgments. As stated in *In re Winship*,²¹² it offers an "assessment of the comparative social disutility" of trial errors.²¹³ Blackstone's adage that "it is better that ten guilty persons escape, than that one innocent suffer,"²¹⁴ is a cornerstone in legal doctrine. It is consistent with decision theoretic analyses,²¹⁵ as well as with assessments observed in empirical studies.²¹⁶ Lowering the standard expresses an acceptance of a higher ratio of erroneous convictions. But there is no reason for such an acceptance.

²¹⁰ See, for example, Jeffrey J. Rachlinski, *Heuristics and Biases in the Courts: Ignorance or Adaptation?*, 79 Or L Rev 61 (2000) (highlighting law's adaptive characteristic). On the efficiency of the common law, see Richard A. Posner, *Economic Analysis of Law* Part 2 (Aspen 6th ed 2003).

²¹¹ The reason for focusing on the standard of proof is that it is more intimately connected with the rate of conviction than evidentiary rules and other trial procedures, and it is also readily modifiable.

²¹² 397 US 358 (1970).

²¹³ Id at 371 (Harlan concurring).

²¹⁴ William M. Blackstone, 4 *Commentaries* at *358. For a helpful discussion, see Lawrence Solan, *Refocusing the Burden of Proof in Criminal Cases: Some Doubt about Reasonable Doubt*, 78 Tex L Rev 105 (1999) (discussing the history and motivations behind the standard).

²¹⁵ The value of 90 percent shows up in both normative and descriptive analyses. According to decision theory, a ratio of nine false acquittals to one false conviction should lead to a standard of proof of about 90 percent. See Francis C. Dane, *In Search of Reasonable Doubt: A Systematic Examination of Selected Quantification Approaches*, 9 L & Human Behav 141, 154–56 (1985) (testing several quantification approaches to determine practical reasonable doubt). But see Erik Lillquist, *Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability*, 36 UC Davis L Rev 85, 114–17 (2002) (criticizing these conclusions and suggesting supplementing the calculus with the utilities of correct acquittals and correct convictions).

²¹⁶ Research that monitors people's perceptions of the standard reveals similar values. While assessments vary somewhat with the methodology of elicitation, the responses tend to converge in the range of 80 percent to 90 percent. See Reid Hastie, *Algebraic Models* at 102 (cited in note 147) (listing results from numerous empirical studies).

Another objection is that a moderating intervention is unnecessary because its function is already taken care of by the adversarial nature of the trial; even when the arguments of defense attorneys fail to sway jurors, they might nonetheless moderate the coherence effect. Although advocates' arguments may conceivably have a moderating effect, the research suggests that it will be a limited one. Recall that in both the *Quest* and *Jason Wells* cases, participants were presented with the arguments by both lawyers before being asked to decide the verdict. Yet, we observed substantial coherence shifts. Research on persuasion also shows that the effectiveness of a persuasive message depends upon the target's perception of the source. Most notably, persuasion is adversely affected when the source is deemed to lack credibility.²¹⁷ Thus, a juror is not likely to respond to the urging of an advocate for the disbelieved party, especially when the juror is already close to making up her mind. Moreover, given the lopsidedness of the juror's representation, the opposing advocate's arguments might appear to be extreme, and thus become even less credible.²¹⁸ There is also a danger that the arguments will backfire; in some situations, resistance to persuasion can entrench one's position.²¹⁹

Another objection is that the jury instructions given at trial render a moderating procedure unnecessary. Jury instructions contain a barrage of directives and admonitions intended to prevent error and bias in verdict decisions. For example, judges tell jurors to presume innocence, to base their decisions on all the relevant evidence, to give careful attention to the testimony and exhibits, and to keep an open

²¹⁷ The consequence of source credibility on the effectiveness of persuasion is broadly acknowledged, and has become a classic experimental finding. See, for example, Shelly Chaiken and Durairaj Maheswaran, *Heuristic Processing Can Bias Systematic Processing: Effects of Source Credibility, Argument Ambiguity, and Task Importance on Attitude Judgment*, 66 *J Personality & Soc Psych* 460, 464–66 (1994); Martin Fishbein and Icek Ajzen, *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research* 451–509 (Addison-Wesley 1975); Carl I. Hovland, Irving L. Janis, and Harold H. Kelley, *Communication and Persuasion: Psychological Studies of Opinion Change* 25–48 (Yale 1953).

²¹⁸ The “accentuation effect” is the tendency to exaggerate the discrepancies between one's own position and those expressed by people on the opposite side of the issue. The stronger one's own views, the more extreme the other position seems. See Robyn M. Dawes, David Singer, and Frank Lemons, *An Experimental Analysis of the Contrast Effect and Its Implications for Intergroup Communication and the Indirect Assessment of Attitude*, 21 *J Personality & Soc Psych* 281, 294 (1972) (“Subjects in the [] experiments had a strong tendency to exaggerate the differences in attitude between themselves and the people with whom they disagree.”); Robert J. Robinson, et al, *Actual versus Assumed Differences in Construal: “Naive Realism” in Intergroup Perception and Conflict*, 68 *J Personality & Soc Psych* 404 (1995) (studying differences of opinion on abortion and race).

²¹⁹ See Zakary L. Tormala and Richard E. Petty, *What Doesn't Kill Me Makes Me Stronger: The Effects of Resisting Persuasion on Attitude Certainty*, 83 *J Personality & Soc Psych* 1298 (2002) (showing that high certainty in one's initial attitudes increased resistance to subsequent attacks and further enhanced the correspondence between attitudes and behavioral intentions).

mind until the entire case has been submitted for deliberation.²²⁰ However desirable, these instructions are unlikely to have an appreciable impact on coherence shifts. The debiasing experiment indicates that coherence effects are moderated only when participants are specifically instructed to actively generate an alternative mental representation.²²¹ Judicial instructions contain no such directives.

Even if we were to introduce a moderating instruction, its effectiveness would be unclear. There is indeed doubt whether jury instructions produce their intended effects. Research has shown that many instructions are not properly comprehended by jurors,²²² while others demand that jurors perform tasks that are cognitively very difficult, if not outright impossible.²²³ Neither of these difficulties applies to the moderating intervention I have proposed; the instruction is clear and readily feasible. The effectiveness of the proposed instruction might also be hindered by limits of human attention. Jury instructions are already very long, they are often communicated poorly, and there is little doubt that jurors cannot apply or even recall them in full.²²⁴ Thus, it is quite possible that an additional instruction will simply fall by the wayside. Therefore, I suggest that judges take simple steps to increase the instruction's effectiveness, including repeating it, placing it at the beginning and end, resorting to emphatic language, and even requiring jurors to write down reasons supporting the opposite conclusion.

In sum, the adverse effects of coherence shifts on criminal verdicts merit consideration of a moderating instruction.²²⁵

²²⁰ See, for example, O'Malley, Grenig, and Lee, *Federal Jury Practice & Instructions* § 10.01 (cited in note 151) (listing suggested jury instructions); *California Jury Instructions: Criminal* § 0.50 at 4–9 (West 6th ed 1996 & Supp 2002) (providing a preliminary admonition to the jury).

²²¹ See Part I.B.8.

²²² See, for example, Darryl K. Brown, *Regulating Decision Effects of Legally Sufficient Jury Instructions*, 73 S Cal L Rev 1105, 1105 (2000) (indicating cases and empirical studies suggesting that juries often get instructions wrong); Stephen P. Garvey, Sheri Lynn Johnson, and Paul Marcus, *Correcting Deadly Confusion: Responding to Jury Inquiries in Capital Cases*, 85 Cornell L Rev 627, 628–33 (2000) (discussing jury misunderstandings). For a review of the empirical research, see Phoebe C. Ellsworth and Alan Reifman, *Juror Comprehension and Public Policy: Perceived Problems and Proposed Solutions*, 6 Psych, Pub Pol & L 788, 788–92 (2000).

²²³ See, for example, Lempert, 1 Intl J Evid & Proof at 316–17 (cited in note 168). For a review of the empirical research, see Joel D. Lieberman and Bruce D. Sales, *What Social Science Teaches Us about the Jury Instruction Process*, 3 Psych, Pub Pol & L 589 (1997).

²²⁴ See, for example, Peter Tiersma, *The Rocky Road to Legal Reform: Improving the Language of Jury Instructions*, 66 Brooklyn L Rev 1081, 1084–85 (2001) (noting studies that conclude that instructions are hard for juries to remember, understand, and apply). See also Lieberman and Sales, 3 Psych, Pub Pol & L at 623–39 (cited in note 223) (recounting various empirical studies on jury memory and comprehension).

²²⁵ This reform could benefit from further critical examination and debate. In addition to replicating the moderating effect, it would be worth verifying that it will not have a debilitating effect on jurors' ability to decide on a verdict. There is good reason to believe that the moderating intervention will improve group deliberation, but it might also have other unintended effects. I am currently involved in experimentation designed to replicate the effect and to evaluate its

D. Reducing Error in Harmless-Error Analysis

Legal decision-makers rarely have a choice as to the nature of their tasks. One notable exception is appellate review of harmless error, a doctrine that has been the subject of a jurisprudential split between two modes of analysis. One approach focuses on the impact of the error on the trial outcome, and the other hinges on assessing guilt as it appears from the remaining, presumably untainted, evidence. Coherence-based reasoning research indicates that the latter is more prone to erroneous judgments.

Harmless error is one of the most frequently discussed issues in criminal appeals.²²⁶ The power of appellate courts to set aside lower court verdicts is of crucial importance to the administration of justice.²²⁷ The harmless-error doctrine is a direct and powerful means by which appellate courts regulate and censure trial judges, prosecutors, and police investigators. Though it applies similarly to criminal and civil law, the following discussion will focus on the former, where it is more frequently litigated and more hotly contested.

While the right to a fair trial is a fundamental constitutional right, the ideal, error-free trial proves to be rather elusive in practice.²²⁸ The conflicting interests derive from Herbert Packer's dichotomy between the Crime Control and Due Process models of the criminal process.²²⁹ One obviously crucial objective of the trial is to make accurate factual determinations as to the question of the defendant's guilt or innocence.²³⁰ Naturally, this objective is frustrated when true determinations of guilt are reversed due to inconsequential investigative or trial flaws.²³¹ Yet another principal objective of the trial is to enforce the

advantages and potential collateral effects.

²²⁶ See William M. Landes and Richard A. Posner, *Harmless Error*, 30 J Legal Stud 161, 161 (2001).

²²⁷ As Wigmore explained, whether to grant a new trial is "a great question, because . . . the whole status of the law of evidence, as well as the efficiency of our methods of doing justice, is dependent upon the answer." John H. Wigmore, *New Trials for Erroneous Rulings upon Evidence: A Practical Problem for American Justice*, 3 Colum L Rev 433, 433 (1903). See also Roger J. Traynor, *The Riddle of Harmless Error* 80 (Ohio State 1970) ("[T]he evaluation of an error as harmless or prejudicial is one of the most significant tasks of an appellate court, as well as one of the most complex.").

²²⁸ The Court repeatedly reminds us that "the Constitution entitles a criminal defendant to a fair trial, not a perfect one." *Delaware v Van Arsdall*, 475 US 673, 681 (1985).

²²⁹ See Herbert L. Packer, *The Limits of the Criminal Sanction* 149–73 (Stanford 1968).

²³⁰ See, for example, *Van Arsdall*, 475 US at 681 ("The harmless-error doctrine recognizes the principle that the central purpose of a criminal trial is to decide the factual question of the defendant's guilt or innocence.").

²³¹ See, for example, *Kotteakos v United States*, 328 US 750, 760 (1946) (concluding that the congressional intent in limiting the scope for reversal was to preserve the right to a fair trial "without giving men fairly convicted the multiplicity of loopholes which any highly rigid and minutely detailed scheme of errors, especially in relation to procedure, would engender and reflect in a printed record").

safeguards owed to criminal defendants.²³² Letting convictions stand despite failings of the officers of the criminal justice system increases trial error and diminishes the judiciary's ability to deter the recurrence of such flaws.²³³ In the long run, consistent rebukes by appellate courts would have a cleansing effect that both protects defendants' rights and results in an increase in trial accuracy.²³⁴ Both sides to the debate claim correctly that inattention to their respective objective is bound to erode the public's confidence in the criminal justice system.²³⁵

Needless to say, these objectives are not amenable to an easy balancing judgment, and the terse language of the Federal Rules of Criminal Procedure offers little guidance.²³⁶ Devising the appropriate test for harmless-error analysis has been the topic of much jurisprudential contention.²³⁷ In grappling with this task, courts have analyzed

Harmless-error doctrine developed as a reaction to a state in which federal courts were viewed as "impregnable citadels of technicality." Marcus A. Kavanagh, *Improvement of Administration of Criminal Justice by Exercise of Judicial Power*, 11 ABA J 217, 222–23 (1925).

²³² See, for example, *Rose v Clark*, 478 US 570, 588 (1986) (Stevens concurring) ("[O]ur Constitution, and our criminal justice system, protect other values besides the reliability of the guilt or innocence determination. A coherent harmless-error jurisprudence should similarly respect those values.").

²³³ See, for example, *id.* at 588–89 (warning that a broad application of the harmless-error rule "can only encourage prosecutors to subordinate the interest in respecting the Constitution to the ever-present and always powerful interest in obtaining a conviction in a particular case"); *Harrington v California*, 395 US 250, 255 (1969) (Brennan dissenting) (warning that the deterrent effect of constitutional protections "on the actions of both police and prosecutors, not to speak of trial courts, will be significantly undermined").

²³⁴ Justice Traynor argued, "If appellate judges forthrightly opened the way to a new trial whenever a judgment was contaminated by error, there would be a cleansing effect on the trial process. A sharp appellate watch would in the long run deter error at the outset, thereby lessening the need of appeal and retrials." Traynor, *Riddle of Harmless Error* at 50 (cited in note 227).

²³⁵ Compare *Van Arsdall*, 475 US at 681 (noting the principle that the fact that trials are intended foremost to decide the factual question of guilt or innocence "promotes public respect for the criminal process by focusing on the underlying fairness of the trial rather than on the virtually inevitable presence of immaterial error"), with *Rose*, 478 US at 588 (Stevens concurring) (emphasizing that protecting constitutional values besides the reliability of the determination of guilt or innocence serves to strengthen public confidence in the administration of justice).

²³⁶ The determination of harmfulness hinges on whether the error affected the defendant's "substantial rights." See FRCrP 52(a) ("[Any] error, defect, irregularity, or variance that does not affect substantial rights must be disregarded."). See also FRCP 61 ("The court at every stage of the proceeding must disregard any error or defect in the proceeding which does not affect the substantial rights of the parties.").

²³⁷ For representative scholarly discussions, see generally Sam Kamin, *Harmless Error and the Rights/Remedies Split*, 88 Va L Rev 1, 72, 78–86 (2002) (arguing that the state should be precluded from claiming harmless error when they knew or should have known they were violating the defendant's constitutional rights); Landes and Posner, 30 J Legal Stud at 162–69 (cited in note 226) (establishing an economic model to maximize social welfare in harmless-error rulings); Harry T. Edwards, *To Err Is Human, But Not Always Harmless: When Should Legal Error Be Tolerated?*, 70 NYU L Rev 1167, 1205–09 (1995) (arguing tentatively for a bright-line approach to the harmless-error analysis of trial errors implicating the constitutional rights of the defendant); Gregory Mitchell, *Against "Overwhelming" Appellate Activism: Constraining Harmless Error Review*, 82 Cal L Rev 1335, 1341 (1994) (arguing that "the overwhelming-evidence approach to harmless error review is much too permissive of appellate activism, whereas the contribution-

cases along various dimensions, including the constitutional nature of the error,²³⁸ the degree of certainty of harmfulness,²³⁹ and since *Arizona v. Fulminante*,²⁴⁰ the classification of the error as “structural” or mere “trial error.”²⁴¹

Most of the judicial and scholarly attention has focused on yet another dimension—whether the harm should be determined based on the effect of the error itself or on the remainder of the evidence.²⁴² According to the rule announced in *Chapman v. California*,²⁴³ the harmless-error test hinges on the reviewing judge’s assessment of the impact the error had on the jury’s verdict.²⁴⁴ In contrast to this *error-focused* approach,²⁴⁵ *Harrington v. California*²⁴⁶ proposes that the reviewing judge should assess whether the conviction would still have resulted in the absence of error.²⁴⁷ According to this *guilt-focused* ap-

to-conviction approach preserves the appellate court’s well-justified deferential role regarding findings of guilt”); Henry P. Monaghan, *Harmless Error and the Valid Rule Requirement*, 1989 S Ct Rev 195, 196 (arguing that harmless-error analysis is not appropriate if a state court conviction is based on a constitutionally infirm rule); Tom Stacy and Kim Dayton, *Rethinking Harmless Constitutional Error*, 88 Colum L Rev 79, 91–98 (1988) (grounding the harmless-error review in the purpose of the right potentially infringed upon); Martha A. Field, *Assessing the Harmlessness of Federal Constitutional Error—A Process in Need of a Rationale*, 125 U Pa L Rev 15, 60–61 (1976) (advocating an evaluation of the evidence to be excluded to determine its persuasiveness); Stephen A. Saltzburg, *The Harm of Harmless Error*, 59 Va L Rev 988, 989–98 (1973) (advocating for different harmless-error review standards in civil and criminal contexts); Traynor, *Riddle of Harmless Error* at 33–37 (cited in note 227) (suggesting the application of a “high probability [of] harmlessness” test to help evaluate which errors are, in fact, harmless).

²³⁸ Prior to 1967, only nonconstitutional errors were subject to harmless-error review. That changed with the landmark case of *Chapman v. California*, 386 US 18, 22 (1967).

²³⁹ See *Brecht v. Abrahamson*, 507 US 619, 637 (1993) (shifting the burden to defendants who will not be entitled to habeas relief “unless they can establish that the trial error resulted in ‘actual prejudice’”). For a critical commentary, see John H. Blume and Stephen P. Garvey, *Harmless Error in Federal Habeas Corpus after Brecht v. Abrahamson*, 35 Wm & Mary L Rev 163 (1993).

²⁴⁰ 499 US 279 (1991).

²⁴¹ *Id.* at 307–12 (differentiating trial error from structural errors for the purposes of harmless-error analysis). For scholarly reactions to the dichotomy, see Charles J. Ogletree, Jr., *Arizona v. Fulminante: The Harm of Applying Harmless Error to Coerced Confessions*, 105 Harv L Rev 152, 161–65 (1991) (arguing that the Court’s distinction between trial and structural errors is unpersuasive); David McCord, *The “Trial”/“Structural” Error Dichotomy: Erroneous, and Not Harmless*, 45 Kan L Rev 1401, 1454 (1997) (arguing that the dichotomy is confusing to lower courts, further complicating the harmless-error analysis).

²⁴² See note 237.

²⁴³ 386 US 18 (1967).

²⁴⁴ *Id.* at 20.

²⁴⁵ Judge Edwards labels this approach the *effect-on-the-verdict* approach. Edwards, 70 NYU L Rev at 1171 (cited in note 237).

²⁴⁶ 395 US 250 (1969).

²⁴⁷ *Id.* at 254 (concluding that the overwhelming guilt established by admissible evidence negates the trial error). See also *Brown v. United States*, 411 US 223, 231 (1973) (citing the *Harrington* Court’s harmless-error standard); *Schneble v. Florida*, 405 US 427, 430–31 (1972) (comparing *Schneble* to *Harrington*); *Milton v. Wainwright*, 407 US 371, 377 (1972) (holding that the jury would have reached the same verdict even if the defendant’s statements to the police had not been admitted into evidence). In *Rose v. Clark*, for example, the Court stated: “Where a re-

proach, the error itself plays a marginal role; the analysis hinges rather on the strength of the remainder of the evidence.²⁴⁸ As long as the unchallenged evidence provides solid grounds for conviction, even grave errors will be deemed harmless.

Over the past thirty-five years, the Burger and Rehnquist Courts have gradually expanded the doctrine of harmless error,²⁴⁹ with fewer flawed trials reversed.²⁵⁰ A brief flirtation with the error-focused approach in *Sullivan v Louisiana*²⁵¹ aside, the Court usually relies on the *Harrington* guilt-focused test.²⁵² Legal scholarship, on the other hand, has generally repudiated the guilt-focused approach. Critics contend that reviewing the entire record on appeal undermines the right to a trial by jury, is unnecessarily speculative, and is based on incomplete accounts of the evidence.²⁵³

By itself, coherence-based reasoning research cannot resolve the jurisprudential dispute across the entire spectrum of harmless-error

viewing court can find that the record developed at trial establishes guilt beyond a reasonable doubt, the interest in fairness has been satisfied and the judgment should be affirmed.” 478 US at 579.

²⁴⁸ As Justice Traynor explains, the “correct result” test asks whether the jury reached the correct conclusion notwithstanding the admission of tainted evidence. Traynor, *Riddle of Harmless Error* at 18–22 (cited in note 227). See also Edwards, 70 NYU L Rev at 1171 (cited in note 237) (arguing that a judgment of whether error is harmless is actually a judgment of how the appellate judge would have decided guilt). Henry Monaghan has distinguished the two approaches as “jury-centered” and “judge-centered.” Monaghan, 1989 S Ct Rev at 203 (cited in note 237).

²⁴⁹ “[I]t is the rare case in which a constitutional violation will not be subject to harmless-error analysis.” *Sullivan v Louisiana*, 508 US 275, 282 (1993) (Rehnquist concurring).

²⁵⁰ See Edwards, 70 NYU L Rev at 1180–81 (cited in note 237) (citing numbers suggesting that applications of the harmless-error doctrine increased significantly post-*Chapman*); Mitchell, 82 Cal L Rev at 1348 n 82 (cited in note 237) (suggesting that guilt-focused approach review is most commonly used in harmless-error determinations).

²⁵¹ 508 US 275, 279 (1993) (“The inquiry . . . is not whether, in a trial that occurred without the error, a guilty verdict would surely have been rendered, but whether the guilty verdict actually rendered in *this* trial was surely unattributable to error.”).

²⁵² See, for example, cases in note 247. See also *Neder v United States*, 527 US 1, 17 (1999) (concluding beyond a reasonable doubt that the jury verdict would have been the same absent the error).

²⁵³ See, for example, Kamin, 88 Va L Rev at 72–73 (cited in note 237) (arguing that the guilt-focused harmless-error approach allows courts to avoid answering the substantive questions of constitutional law around which the errors are made); Linda E. Carter, *The Sporting Approach to Harmless Error in Criminal Cases: The Supreme Court’s “No Harm, No Foul” Debacle in Neder v. United States*, 28 Am J Crim L 229, 239–45 (2001) (claiming that a guilt-focused harmless-error review unconstitutionally limits the right to trial by jury); Edwards, 70 NYU L Rev at 1192–94 (cited in note 237) (arguing that review in a guilt-focused manner denies the right to trial by jury); Mitchell, 82 Cal L Rev at 1353–57 (cited in note 237) (arguing that appellate review is a poor place to evaluate the record, undermining the guilt-focused harmless-error review); Traynor, *Riddle of Harmless Error* at 25–33 (cited in note 227) (raising multiple arguments against the appellate review of harmless error with the guilt-focused approach including difficulty with burdens, task confusion, and so on). See also Charles S. Chapel, *The Irony of Harmless Error*, 51 Okla L Rev 501, 506 (1998) (describing the guilt-focused approach as inimical to constitutional protections); Erwin Chemerinsky, *No Harm, No Foul*, 16 Cal Law 27 (Jan 1996) (criticizing the California Supreme Court’s use of the guilt-focused approach).

cases. It can, however, offer a strong reason for preferring the error-focused approach in cases that involve an erroneous introduction of inadmissible evidence at trial, such as a coerced confession, hearsay, prejudicial evidence, or a *Bruton* violation.²⁵⁴ In these cases, the reviewing judge is of course privy to the content of the contested evidence. Coherence-based reasoning presents the possibility that having been exposed to the entire record, judges might be influenced by the impermissible evidence. As discussed above, coherence shifts are effectuated via intricate interconnections, so that variables can influence other variables to which they are not directly related. Sufficiently strong pieces of evidence can affect the entire mental model of the case, indirectly changing the perceptions of unrelated pieces of evidence.²⁵⁵ Since coherence effects occur without awareness,²⁵⁶ the judge will decide in accordance with her perception of the evidence, which, unbeknownst to her, has likely been skewed by the illicit variable.²⁵⁷ Thus, the guilt-focused approach is undermined by the danger that the supposedly untainted remainder of the evidence is anything but untainted. In criminal trials, where inadmissible evidence is invariably prosecution evidence, guilt-focused analysis will tend to lead to dubious affirmations of convictions.

This concern is strongest in cases where there are independent reasons for doubting the underlying veracity of the other evidence. Take for example the case of *Milton v Wainwright*,²⁵⁸ where the Court held harmless an impermissible surreptitious interrogation in violation of a defendant's *Massiah* right.²⁵⁹ Basing its decision on an extensive examination of the trial record, the Court found what it considered to be overwhelming evidence of guilt, most notably three confessions made by the defendant prior to the indictment. Those confessions, however, were themselves of questionable reliability because they were obtained by dubious means. Indeed, as the dissent pointed out, the likely reason for conducting the surreptitious investigation was that those confessions might not have held up in court.²⁶⁰ The bootstrapping in this case is apparent: the illicit interrogation was held harmless based on the strength of the confessions, while it is quite pos-

²⁵⁴ See *Bruton v United States*, 391 US 123, 137 (1968) (excluding the statement of a non-testifying codefendant).

²⁵⁵ See Part I.B.5.

²⁵⁶ See Part I.B.2.

²⁵⁷ This danger has been pointed out by a sitting judge. See Edwards, 70 NYU L Rev at 1205 (cited in note 237) (noting that once the judge has seen the evidence, he faces the risk of being influenced by it).

²⁵⁸ 407 US 371 (1972).

²⁵⁹ *Id.* at 372. In *Massiah v United States*, 377 US 201 (1964), the Court found a constitutional violation in surreptitiously interrogating a defendant out of the presence of his lawyer. *Id.* at 207.

²⁶⁰ *Milton*, 407 US at 383 (Stewart dissenting).

sible that the veracity of the confessions was bolstered by the information obtained in the illicit interrogation.²⁶¹

The practical differences between the two analytical approaches are unlikely to be as stark as implied by the doctrinal debate. In all cases, the judge is aware of the evidence in question, and there is ample experimental evidence showing that people simply cannot shut off knowledge at will.²⁶² Quite possibly, error-focused analyses are not completely immune from contamination: due to coherence effects, the impermissible evidence can make the error seem less egregious. However, the narrower the judgment, the less susceptible it is to coherence effects. The fewer the dimensions of the decision, the more likely that the choice will be based on the decision-maker's evaluations and beliefs that precede, and are thus exogenous to, the decision itself. Furthermore, unlike assessments of the entire record, the harmfulness of errors is relatively commensurable and classifiable according to a set of rules and standards.²⁶³ Precursors of such a scheme can be discerned from some Supreme Court decisions.²⁶⁴

Harmless-error analysis pertains also to cases that do not involve admissibility of evidence, such as when the trial contained a faulty jury instruction, an improper comment about the defendant's failure to testify, or a violation of the defendant's Confrontation Clause rights. In the absence of impermissible evidence, there is no direct danger of contamination of the remainder of the evidence through indirect coherence effects. Therefore, coherence-based reasoning does not offer a strong prescription. It does, nonetheless, provide some unique insights into the jurisprudential dilemma by adding weight to the error-based approach. These observations apply equally to evidentiary and non-evidentiary errors.

First, recall that appellate reversals serve important constitutional functions by condemning the infringement of the defendant's rights; educating police investigators, prosecutors, and trial judges; and deterring them from future violations. Coherence-based reasoning highlights the extent to which these functions are compromised when errors are declared harmless. A basic finding of the research is that the

²⁶¹ Id. For a similar example, see *Schneble*, 405 US at 434–36 (Marshall dissenting) (criticizing the Court's decision to hold harmless the admission of a statement of a nontestifying codefendant in light of the defendant's confession despite evidence of police coercion).

²⁶² For a comprehensive review of people's limited ability to ignore information, see generally Jonathan M. Golding and Colin M. MacLeod, eds, *Intentional Forgetting: Interdisciplinary Approaches* (Lawrence Erlbaum 1998).

²⁶³ For example, some violations are bound to be classified as never warranting harmless-error analysis, as is currently the case with respect to deprivation of the right to counsel and the right to an impartial judge. See *Chapman*, 386 US at 43–44 (Stewart concurring). Most others can be graded according to the assessed severity of their impact on a jury decision.

²⁶⁴ See note 275.

variables that support the defeated alternative are dismissed, rejected, or ignored. In other words, the losing values and principles are devalued.

In some cases, the devaluation is subtle, reflected only in the tone of the opinion, failure to mention the importance of the right denied by the error, selectivity in sources cited, and the like.²⁶⁵ In other cases, however, the devaluation can be quite glaring, such as in *Darden v Wainwright*,²⁶⁶ where the Court marginalized the conduct of a prosecutor who repeatedly expressed his wish to see the defendant get “blown away by a shotgun,”²⁶⁷ and in *Arizona v Fulminante*, where, relying on a questionable assertion about the impact of involuntary confessions, the Court trivialized the legal significance of having them admitted into the record.²⁶⁸ The Court has even gone so far as to devalue the significance of the harmless-error doctrine itself.²⁶⁹

²⁶⁵ Note, for example, the selective citation of Roger Traynor’s classic text, *The Riddle of Harmless Error* (cited in note 227). His quote “Reversal for error, regardless of its effect on the judgment, encourages litigants to abuse the judicial process and bestirs the public to ridicule it,” id at 50, has been cited exclusively in opinions that find the errors harmless. See, for example, *Neder*, 527 US at 18; *Johnson v United States*, 520 US 461, 470 (1997); *Van Arsdall*, 475 US at 681; *Rose*, 478 US at 577. On the other hand, the quotation “In the long run there would be a closer guard against error at the trial, if . . . courts were alert to reverse, in case of doubt, for error that could have contaminated the judgment,” Traynor, *Riddle of Harmless Error* at 23, has been cited only in a case in which the conviction was ordered reversed, *O’Neal v McAninch*, 513 US 432, 442 (1995).

²⁶⁶ 477 US 168 (1986).

²⁶⁷ Id at 180 n 12.

²⁶⁸ Chief Justice Rehnquist stated:

The evidentiary impact of an involuntary confession, and its effect upon the composition of the record, is indistinguishable from that of a confession obtained in violation of the Sixth Amendment—of evidence seized in violation of the Fourth Amendment—or of a prosecutor’s improper comment on a defendant’s silence at trial in violation of the Fifth Amendment.

Fulminante, 499 US at 310. The Chief Justice went on to claim that relief cannot be based on a belief that “there is something more ‘fundamental’ about involuntary confessions.” Id at 311. This part of the opinion was joined by Justices O’Connor, Kennedy, Souter, and Scalia. Justice Kennedy though, cautioned against ignoring the

indelible impact a full confession may have on the trier of fact. . . . If the jury believes that a defendant has admitted the crime, it doubtless will be tempted to rest its decision on that evidence alone, without careful consideration of the other evidence in the case. Apart, perhaps, from a videotape of the crime, one would have difficulty finding evidence more damaging to a criminal defendant’s plea of innocence.

Id at 313 (Kennedy concurring).

Prompted by the Chief Justice’s assertion, legal psychologist Saul Kassin sought to test its empirical validity. Mock jurors were asked to determine guilt in four criminal trials (murder, rape, assault, and theft) in which the prosecutorial evidence included a confession, an eyewitness identification, or testimony about the defendant’s character. It was found that evidence of confessions produced higher conviction rates than the other forms of incriminating evidence, suggesting unique impact of self-incriminating confessions. Saul M. Kassin and Katherine Neumann, *On the Power of Confession Evidence: An Experimental Test of the Fundamental Difference Hypothesis*, 21 L & Human Behav 469 (1997). In a different experiment it was found that confes-

Under the guilt-focused approach, findings of harmfulness are influenced primarily by the strength of the evidence, which is essentially unrelated to the egregiousness of the error. Strong evidence of guilt can trump even grave error and misconduct. This arbitrary relationship between the strength of the evidence and the gravity of constitutional violation leads to an ad hoc jurisprudence that undercuts the expressive, educational, and deterrent functions of appellate review. A doctrine based on error-focused analysis would likely spur a jurisprudence that would better achieve these goals.²⁷⁰ Second, focusing on the evidence of guilt can have the unintended effect of arousing the judge's anger toward the defendant. As Judge Edwards describes, "A wrong, often a grievous wrong, has occurred, and the defendant, by all appearances, is responsible. It is, therefore, to be expected that the desire to punish the guilty will frequently prevail over the need to honor individual rights."²⁷¹ Empirical research suggests that upon witnessing injustice, people display a need to re-establish a sense of justice and social order, especially in light of the possibility that the misdeed will go unpunished. It has been shown that exposure to anger-provoking stimuli increases the tendency to blame other people for ambiguous events and to neglect alternative explanations and possible mitigating circumstances.²⁷² These findings prevail even when the target of the judgment has not been identified as the person who caused the anger-provoking event.²⁷³ The experimental finding that coherence effects

sions have an indelible effect on criminal verdicts rendered by mock jurors. See Saul M. Kassir and Holly Sukeel, *Coerced Confessions and the Jury: An Experimental Test of the "Harmless Error" Rule*, 21 L & Human Behav 27 (1997).

²⁶⁹ Chief Justice Burger explained that "the integrity of the process carries less weight, for it is the essence of the harmless-error doctrine that a judgment may stand only when there is no 'reasonable possibility that the [practice] complained of might have contributed to the conviction.'" *United States v Hasting*, 461 US 499, 506 (1983) (citation omitted).

²⁷⁰ True, there is a danger that when a court finds an error to be harmless based on the error-based approach, the opinion might minimize the magnitude of the error. I believe, however, that this is the lesser evil. First, given the type of analysis, only the less egregious errors would be found to be harmless; egregious errors would be declared harmful regardless of the evidence. Second, in the long run, error-focused analysis will better fulfill the expressive, educative, and deterrent functions and thus reduce the incidence of error.

²⁷¹ Edwards, 70 NYU L Rev at 1194 (cited in note 237). Similarly, Judge Posner has explained that "it takes a highly disciplined judge to vote to reverse a conviction when he thinks the defendant is guilty." Richard A. Posner, *An Economic Approach to the Law of Evidence*, 51 Stan L Rev 1477, 1518 (1999).

²⁷² See Jennifer S. Lerner, Julie H. Goldberg, and Philip E. Tetlock, *Sober Second Thought: The Effects of Accountability, Anger, and Authoritarianism on Attributions of Responsibility*, 24 Personality & Soc Psych Bull 563 (1998) (demonstrating the effects of anger on judgments of blame); Brian M. Quigley and James T. Tedeschi, *Mediating Effects of Blame Attributions on Feelings of Anger*, 22 Personality & Soc Psych Bull 1280, 1283-84 (1996) (same); Dacher Keltner, Phoebe C. Ellsworth, and Kari Edwards, *Beyond Simple Pessimism: Effects of Sadness and Anger on Social Perception*, 64 J Personality & Soc Psych 740, 743-45 (1993) (demonstrating the effect of anger on attributing blame to personal factors).

²⁷³ See Julie H. Goldberg, Jennifer S. Lerner, and Philip E. Tetlock, *Rage and Reason: The*

can be driven by affective states explains how a state of anger can lead to misperceptions of the evidence.²⁷⁴ Dwelling on the often-troubling details of the criminal event can arouse the judge's anger. In close cases, with the defendant somehow implicated in the crime and with no one else to blame, there is a danger that the judge's mental representation of the case will shift toward supporting a conclusion of guilt.

Three final points are in order. First, it should not be expected that error-focused analyses entail automatic reversals of convictions. There are ample examples in which error-focused analysis has led to findings of harmlessness.²⁷⁵ Second, while this discussion has focused on the criminal setting, it is in principle applicable to the civil setting as well. Finally, it might be the case that the type of analysis endorsed by the judge bears no real impact on the actual decision. Judges may simply use the guilt-focus and error-focus terminology as a label for decisions to uphold and reverse trial court verdicts. However, to the extent that stated doctrine influences decisions made, judges would be well advised to accommodate for the impact of coherence-based reasoning on their decisions.

CONCLUSION

Coherence-based reasoning offers a much-needed alternative to the stalemate between Rationalist and Critical models of legal decision making. Unlike those two approaches, this body of experimental research offers a glimpse into the black box, dispelling both the sanguine and disparaging views of legal decision making. Decisions and inferences do not conform to models of rationalism inspired by logical forms of inference making, nor are they based on consciously disingenuous, biased, or backward reasoning. Complex decisions are solved rather by nuanced cognitive processes that progress bidirectionally between premises and facts on the one hand, and conclusions on the other. Ultimately, people make decisions through what appears to be a

Psychology of the Intuitive Prosecutor, 29 Eur J Soc Psych 781 (1999) (finding that people lower their thresholds for attributing harmful intent and recommending harsh punishment when they know societal norms have been violated and believe that the perpetrator escaped punishment). Suggestions to this effect have also been made by Neil Vidmar, *Retribution and Revenge*, in Joseph Sanders and V. Lee Hamilton, eds, *Handbook of Justice Research in Law* 31, 56 (Kluwer 2001) (suggesting some evidence for groups displacing punishment against non-anger provokers).

²⁷⁴ See Part I.B.6. The experimentation focused mostly on positive affective states, but there is good reason to believe that similar effects occur with negative affect. Recall also that in one of the versions of the *Quest* study, a malevolent depiction of the defendant triggered corresponding coherence effects. See note 78 and accompanying text.

²⁷⁵ See, for example, *Greer v Miller*, 483 US 756, 766 (1987) (holding that the prosecutor's questioning of respondent's post-arrest silence followed immediately by a judicial admonition was harmless); *Rushen v Spain*, 464 US 114, 120–21 (1983) (finding that ex parte communication between the trial judge and a juror was harmless).

rational-like choice in which a strong alternative is straightforwardly preferred over its rival. However, this dominance is the product of an unconscious cognitive process that reconstructs and transforms difficult and complex decisions into easy ones by amplifying one alternative and deflating the other. This transformation of the mental model of the decision lies at the heart of the decision-making process.

One objective of this Article has been to elucidate the implications of coherence-based reasoning for four important aspects of the trial. I argued that in each of these domains, the design of the trial is premised on erroneous assumptions about human cognition, and that these failings lead to systematic errors. Crucially, I emphasized that these failings are not immutable. Coherence-based reasoning offers not only a set of critical tools, but constructive ones too. By identifying the cognitive phenomena that lie at the root of the failings, we can devise interventions and introduce procedures that should make the decision-making process better fit the legal ideals it is intended to serve. First, I suggested that giving juries more extensive instructions at the start of the trial will make for more legally accurate decisions. Second, the research contributes to the central debate in evidence law between holistic and atomistic approaches. I argued that the endorsement of holism as a descriptive matter does not warrant prescriptive positions advocated by its proponents. Third, I suggested that juries in criminal trials be given a simple instruction to temper the contribution of coherence effects to the likelihood of conviction. Finally, I suggested reasons that weigh against the Supreme Court's approach to harmless-error analyses.

Another objective of this Article has been to introduce legal scholars to a body of research, which has been published thus far in journals of experimental psychology. Coherence-based reasoning is highly relevant to legal decision making. The applicability of this research to the law extends well beyond the four trial issues discussed here and the issue of appellate judging that I have explored in previous work.²⁷⁶ The theory and its underlying research have already been applied by other legal scholars to explain how fact-finders combine the many goals and values they seek in the determination of a verdict,²⁷⁷ the effects of stereotypes in the context of advocacy,²⁷⁸ the de-

²⁷⁶ For a brief summary of the application to appellate judicial reasoning, see text accompanying note 117.

²⁷⁷ See Jennifer K. Robbennolt, John M. Darley, and Robert J. MacCoun, *Symbolism and Incommensurability in Civil Sanctioning: Decision Makers as Goal Managers*, 68 *Brooklyn L Rev* 1121, 1148–57 (2003) (discussing legal decision making as constraint satisfaction).

²⁷⁸ See Gary Blasi, *Advocacy against the Stereotype: Lessons from Cognitive Social Psychology*, 49 *UCLA L Rev* 1241, 1259–66 (2002) (discussing mechanisms of constraint satisfaction in the context of stereotypes).

velopment and maintenance of cultural beliefs,²⁷⁹ and the choice among competing conceptual systems in constitutional law.²⁸⁰

A few brief examples may suggest the potential scope for future applications. The research sheds light on a longstanding dilemma in the doctrine of collateral estoppel.²⁸¹ Coherence-based reasoning counsels a more extensive use of special verdicts in civil trials,²⁸² and it gives reasons to consider special verdicts even in some types of criminal trials.²⁸³ The research can also shed light on ongoing debates concerning larger jurisprudential issues, such as the use of rules versus standards.²⁸⁴ These applications may lead to different conclusions de-

²⁷⁹ See Dan Kahan and Donald Braman, *Modeling Facts, Culture, and Cognition in the Gun Debate* (draft), online at http://www.cse.unsw.edu.au/~lambert/guns/stimulating_simulations_ver_4.5.pdf (visited Apr 4, 2004).

²⁸⁰ See Michael H. Shapiro, *Choosing Conceptual Systems in Constitutional Law, and Some Other Locations* (draft), online at <http://hal-law.usc.edu/faculty/workshops/documents/shapiro.pdf> (visited Apr 4, 2004).

²⁸¹ Jurisdictions are split as to instances where the first decision was based on two independently sufficient grounds (say, when a court finds against the plaintiff in a negligence suit on the grounds that the defendant owed the plaintiff no duty of care and that the harm was not established). Some jurisdictions bar litigation on both issues. See, for example, *Winters v Lavine*, 574 F2d 46, 67 (2d Cir 1978); *In re Carozza*, 167 BR 331, 338–40 (Bankr ED NY 1994). Following the approach adopted by the Restatement, other jurisdictions deny collateral estoppel effect for either one. See Restatement (Second) of Judgments § 27, comment (i) (1982) (“If a judgment of a court of first instance is based on determinations of two issues, either of which standing independently would be sufficient to support the result, the judgment is not conclusive with respect to either issue standing alone.”); *Halpern v Schwartz*, 426 F2d 102, 105–06 (2d Cir 1970) (finding that a decision made on two grounds is not preclusive to either).

When dealing with multiple grounds for a decision, there is a danger that due to coherence effects, one strong issue can spill over and influence the other issue, thus casting doubt as to whether the latter was “necessarily decided,” as required by collateral estoppel doctrine. See Charles Alan Wright, Arthur R. Miller, and Edward H. Cooper, 18 *Federal Practice and Procedure* § 4421 at 536 (West 2d ed 2002) (“Issue preclusion attaches only to determinations that were necessary to support the judgment entered in the first action.”). Coherence-based reasoning thus supports the more cautious approach adopted by the Restatement.

²⁸² Here too, the fear is that due to coherence effects, one strong issue can spill over and influence the other issue. Special verdicts might alleviate this problem because the more isolated the judgments, the less likely they are to be affected by coherence effects.

²⁸³ In cases where the evidence is overwhelming with respect to some elements of the crime, but there is only scant evidence to establish a specific element, there is a danger that, due to coherence effects, the fact-finder will find the defendant guilty despite the absence of evidence on the particular element. In such cases, correctly administered special verdicts could serve to expose that evidentiary deficiency. It should be noted that resorting to special verdicts is bound to have collateral effects, including a possible restraint on the exercise of jury nullification. For concerns over the effect of special verdicts on the jury’s power to nullify, see Lafave, Israel, and King, 5 *Criminal Procedure* § 24.10(a) at 611–15 (cited in note 148); *United States v Spock*, 416 F2d 165, 181 (1st Cir 1969) (suggesting that special instructions restrict the right of the jury to ignore instructions and decide independently).

²⁸⁴ See generally Eric A. Posner, *Standards, Rules, and Social Norms*, 21 *Harv J L & Pub Pol* 101 (1997) (establishing a relationship between the economic analysis of rules versus standards, the economic approach to social norms, and the rule of law); Kathleen M. Sullivan, *Foreword: The Justices of Rules and Standards*, 106 *Harv L Rev* 22 (1992) (examining the rules/standards debate in the Supreme Court during the 1991 Term); Louis Kaplow, *Rules versus Standards: An Economic Analysis*, 42 *Duke L J* 557 (1992) (using economic analysis to determine circumstances

pending on the institutional setting within which the decision is made, particularly with respect to the institutional possibilities for implementing debiasing procedures.

Additional research applies to a number of specific kinds of legal decision making. The findings from the experiment *Dixon v Providential Insurance* shed new light on the behavior of legal actors in pre-trial negotiations.²⁸⁵ The horse race study briefly mentioned above offers insight into the issue of gambling behavior.²⁸⁶ The findings showed that novices and experts alike tend to distort the probability of winning toward optimistic overconfidence. This means that gambling establishments benefit considerably from a cognitive distortion. These findings should contribute to the debate surrounding legal regulation of gambling, particularly in the context of state lotteries.

With more experimentation underway, the theory harbors potential for providing unique insight to additional legal applications. I hope that the exposure to this burgeoning body of research will encourage legal scholars to incorporate it further into the analysis of legal decision making. Doing so should provide legal discourse with sharper critical tools and more constructive proposals for reform.

in which rules or standards would be preferable).

Coherence effects are mediated by the ambiguity of the variables: the more ambiguous the issue, the greater the coherence shifts. See text accompanying note 113. It follows that legal directives formulated as standards will likely cause stronger shifts than similar precepts promulgated as rules. The research, then, could provide lawmakers with another consideration in choosing one form or the other.

²⁸⁵ For a brief summary of the findings in *Dixon v Providential Insurance*, see text accompanying notes 92–93. These findings replicate and extend the important research by George Loewenstein and his colleagues on self-serving bias. See Babcock, Loewenstein, and Issacharoff, 22 L & Soc Inquiry at 913 (cited in note 104); George Loewenstein, et al, *Self-Serving Assessments of Fairness and Pretrial Bargaining*, 22 J Legal Stud 135 (1993).

²⁸⁶ For a summary of the findings of the horse race studies, see text accompanying note 91.