

The Redux of Cognitive Consistency Theories: Evidence Judgments by Constraint Satisfaction

Dan Simon, Chadwick J. Snow, and Stephen J. Read
University of Southern California

The authors suggest that decisions made from multiple pieces of evidence are performed by mechanisms of parallel constraint satisfaction, which are related to cognitive consistency theories. Such reasoning processes are *bidirectional*—decisions follow from evidence, and evaluations of the evidence shift toward coherence with the emerging decision. Using a factually complex legal case, the authors observed patterns of coherence shifts that persisted even when the distribution of decisions was manipulated (Study 1) and influenced by the participants' attitudes (Study 2). The evaluations of the evidence cohered with the preferred decision even when participants changed their preference (Study 3). Supporting the bidirectionality of reasoning, Study 4 showed that assigning participants to a verdict affected their evaluation of the evidence. Coherence shifts were observed also in related background knowledge. This research suggests that cognitive consistency theories should play a greater role in the understanding of human reasoning and decision making.

Everyday social reasoning and decision making often rely on multitudes of judgments and inferences. Consider Festinger's (1957) familiar example of a person faced with a decision of whether to accept a dinner invitation or to go to a concert. However trivial, the decision is likely to be complex: It could entail an assessment of the likelihood of enjoying the music to be performed, the probability of offending the host by turning down the invitation, and predictions about the social interaction with each of the people around the dinner table. This task, then, consists of making a number of judgments and inferences and of integrating them into a discrete choice. Despite the ubiquity and importance of these kinds of tasks, the reasoning processes underlying them are still not well understood.

The objective of the current research was to shed light on the process of making decisions that are based on multiple judgments and inferences. We propose that the process can be understood as governed by mechanisms of parallel constraint satisfaction (Holyoak & Simon, 1999; Holyoak & Thagard, 1989; McClelland & Rumelhart, 1986; Read & Miller, 1994; Read, Vanman, & Miller, 1997; Thagard, 1989), which share the gestaltian principles of structural dynamics that underlie classic cognitive consistency theories (Markus & Zajonc, 1985). Such an understanding could support an extension of the scope of cognitive consistency theories to larger, semantically rich, and more complex tasks, which had heretofore been out of their reach (Read et al., 1997; Read & Miller, 1994).

As part of this investigation, we explicitly contrast the predictions of such gestaltian or constraint satisfaction theories with

algebraic models for integration of evidence, such as Norman Anderson's (1996) information integration theory (IIT) and Bayes's theorem. Two of the central assumptions of this class of algebraic models are that the pieces of information are evaluated independently of one another and that their evaluation is completely separated from their subsequent integration into a final decision. Together these two assumptions imply that inference making is *unidirectional*, in that it flows exclusively from the evidence to the conclusion.

In contrast, we hope to show that decision making in these kinds of tasks is processed *bidirectionally*—a term we use to describe processes in which the evidence influences the conclusions and, at the same time, the emerging conclusion affects the evaluation of the evidence. Such a finding would be consistent with cognitive consistency and constraint satisfaction models but strongly inconsistent with IIT and Bayesian models.

Another possibility that we examine is whether the decision-making process affects not only the specific evidence relevant to the task but also related beliefs and background knowledge. Constraint satisfaction models raise the possibility that changes in the evaluation of evidence may also affect the associated beliefs and knowledge that are used to interpret and evaluate the evidence. We investigate these questions in the context of a legal decision-making task in which participants render a discrete verdict in a complex trial that involves numerous pieces of probabilistic evidence.

Cognitive Consistency Theories

Cognitive consistency theories, which flourished between the 1940s and 1960s, were one strand of research that attempted to explore processing in tasks involving multiple variables. These theories were based primarily on Fritz Heider's balance theory (Heider, 1946, 1958) and also included cognitive dissonance theory (Brehm & Cohen, 1962; Festinger, 1957), congruity theory (Osgood & Tannenbaum, 1955), symmetry theory (Newcomb, 1953), and a number of neobalance theories (Abelson & Rosen-

Dan Simon, Law School, University of Southern California; Chadwick J. Snow and Stephen J. Read, Department of Psychology, University of Southern California.

We thank Sam Fraidin, Tom Lyon, Ed McCaffery, and Eric Talley for helpful comments.

Correspondence concerning this article should be addressed to Dan Simon, Law School, University of Southern California, University Park, Los Angeles, CA 90089-0071. E-mail: dsimon@law.usc.edu

berg, 1958; Cartwright & Harary, 1956; Rosenberg, 1960). At the heart of cognitive consistency theories is the gestaltian tenet that human cognition is substantially affected by mutual interaction among pieces of psychological knowledge (Asch, 1946b; Wertheimer, 1922/1967a)—an interaction best understood within the framework of structural dynamics (Markus & Zajonc, 1985; Zajonc, 1968). Cognitive consistency theories were animated by four principles of structural dynamics. First, cognitive states are determined holistically rather than elementally. As Fritz Heider (1960) explained, “The properties of these configurations which determine their meaning and their fate are wholequalities. Consonance or simplicity of the structure cannot be derived in an additive way from the properties of the parts” (p. 168). Similarly, Leon Festinger (1957) explained that a state of dissonance “is not anything which exists all by itself. It is a characterization of a relationship between cognitive elements” (p. 279). Cognitive states, then, are determined not only by the cognitive elements themselves but also by the structures formed through their interaction.

Second, structural properties are dynamic. The interrelatedness of the constitutive elements generates forces that determine the configuration of the structure. Some things “go together,” that is, they are related by cohesive vectors, whereas other things tend to disperse. These forces or vectors determine the stability of the structure and endow it with the potential for movement and change (Zajonc, 1983).

Third, the dynamic character of mental processes is such that they tend to settle at states of distinct structural properties, namely, *Prägnanz* (Wertheimer, 1923/1967c), “good figure” (Heider, 1960; see also Markus & Zajonc, 1985), optimum order (Rosenberg & Abelson, 1960), consonance (Festinger, 1957), or equilibrium (Rosenberg & Abelson, 1960; Tannenbaum, 1968a). In these stable states, “all parts of a unit have the same dynamic character (i.e., all are positive, or all are negative), and entities with different dynamic character are segregated from each other” (Heider, 1946, p. 107). Structural forces keep the equilibrated structures together, whereas states of disequilibrium generate configural forces that drift the system back toward internal coherence (Heider, 1958). In other words, cognitive consistency has a homeostatic property (Rosenberg, 1968), and once the system is driven back to a balanced state, the configural forces of change recede (Rosenberg & Abelson, 1960).

Fourth, and most pertinent to the current experimental project, these dynamic changes that occur at the structural level involve changes, or “reconstructions” (Rosenberg & Abelson, 1960) of the cognitive elements. According to Max Wertheimer (1924/1967b), a fundamental gestaltian principle is that “the part-processes are themselves determined by the intrinsic nature of the whole” (p. 2). In a similar vein, Solomon Asch (1940) offered the insightful distinction between changes in “the judgment of the object” and changes in “the object of judgment” (p. 455). Structural adjustments are brought about by changing the representation of the variables involved. For example, a principal observation of cognitive dissonance theory is that consonance is restored by changing the elements that are in dissonant relations, adding consonant ones, or decreasing the importance of the dissonant elements (Festinger, 1957). Similarly, Asch stated that the tendency to arrive at a consistent, unified view is achieved by getting rid of incompatible elements, if not by some means of an objective examination then

by “distorting the state of affairs” (Asch, 1940, p. 454). Other means of reconstruction include differentiation (Abelson & Rosenberg, 1958), bolstering (Kelman & Baron, 1968), and denial (Rosenberg, 1960). This fourth aspect is of crucial importance to the proposed understanding of human cognition: Reasoning tasks that require the integration of multiple pieces of evidence into a global judgment entail not only making inferences from the evidence to the conclusion but also reverse effects, by which the structural forces that push the system toward good form impose changes on the values of the evidence itself. Hence, the dynamic character of cognitive consistency theories can be characterized as operating in a bidirectional manner—from evidence to conclusions and from conclusions back to evidence (Holyoak & Simon, 1999; Read et al., 1997; Simon & Holyoak, 2002).

Cognitive consistency theories were developed amidst an aspiration toward a general theory of cognition. Kurt Lewin (1935) and other Gestalt psychologists sought to develop a “psychology of knowing” (see Rosenberg & Abelson, 1960). Likewise, William McGuire (1968) had hoped to construct “a psychology of inference,” a framework for the “depiction of the manner and extent to which one idea leads to another psychologically” (p. 140). The endeavor, however, was hobbled by a number of limiting factors. Consistency theories were limited by their inability to represent rich and complex phenomena and by the lack of a formal model for evaluating and computing consistency (Read et al., 1997; Read & Miller, 1994). Consistency theories tended to be designed for relatively small structures; ironically, cognitive dissonance’s explicit confinement to only two elements (Festinger, 1957, p. 13) precluded it from having any bearing on tasks like the dilemma that Festinger posited between going to dinner or to a concert. The theories were confined also to structures in which the strength of the units and the weight of the relations were fixed at unitary levels (e.g., Heider, 1946) and in which the relationships between the units were necessarily “obverse” to each other (Festinger, 1957, p. 13). It is not surprising that with the exception of the thriving research in the forced-compliance paradigm of cognitive dissonance theory (for reviews, see Harmon-Jones & Mills, 1999; Wicklund & Brehm, 1976) and sporadic work on neobalance theories (e.g., Golob, 1974; Insko, 1984), research in cognitive consistency theories came to a virtual standstill by the 1960s.

Algebraic Models and the Story Model

The waning of cognitive consistency theories coincided with the emergence of a very different type of theory. Anderson’s (1962, 1968) IIT is a broad theory of cognition that concentrates on combining psychological stimuli into unitary responses. The conceptual model of IIT contains three sequential operators: valuation, integration, and response. In the valuation phase, physical stimuli or pieces of evidence are psychologically represented in some quantifiable form. The integration phase combines the psychological valuations into an overall judgment. In the response phase, the overall judgment is transformed into an observable response. At the core of IIT is the tenet that human cognition obeys simple algebraic rules, in that judgments of complex phenomena are the mathematical product—typically, a weighted average—of the respective psychological valuations (for reviews, see Anderson, 1981, 1996).

Anderson's cognitive algebra is premised on two central syntactic rules. The first is *meaning invariance*, which stands for the notion that each piece of evidence is evaluated on its own terms and is not affected by the other pieces of evidence (unless there is a preexisting relationship of interdependency). The second rule, *valuation–integration independence*, posits complete separation between the processes of evaluation and integration. In other words, the evaluation of a piece of evidence is assessed independently from how it is combined to form the ultimate conclusion. Together, this pair of rules captures what we describe as the property of unidirectionality: Inferences flow from the individual pieces of evidence toward a computed judgment, but the evaluation of the evidence is in no way affected by the emerging conclusion (valuation–integration independence), and there are no interactions among the pieces of evidence (meaning invariance).

Another broadly accepted theoretical model for the making of probabilistic decisions is Bayes's theorem (for a review, see J. Baron, 2000). At the core of Bayes's theorem is the algebraic exercise of sequential multiplication of the probabilistic values of the event's constitutive elements. Decisions are determined by comparing the posterior likelihood produced by the Bayesian computation with the numerical value assigned to the respective standard of proof (Hastie, 1993). Bayes's theorem relies crucially on the same principles of meaning invariance and evaluation–integration independence. Bayes's theorem does not accommodate any changes in the elements: The numerically expressed probability assigned to each piece of evidence should not be affected by the probabilities of other pieces of evidence or by the computed product. A most natural application for IIT and Bayes's theorem is courtroom decision making. The evaluation of multiple pieces of evidence and their integration into discrete decisions lies at the heart of the legal trial. Indeed, in most contested trials, the evidence is voluminous, fragmentary, incomplete, ambiguous, and inconsistent, and the decision maker must nonetheless produce a categorical verdict of either guilt or innocence. Both IIT (e.g., Kaplan, 1975; Kaplan & Miller, 1978; Ostrom, Werner, & Saks, 1978) and Bayes's theorem (e.g., Friedman, 1997; Kaye, 2000; Lempert, 1977; Schum, 1994) have been applied extensively to courtroom decision making (for a review, see Hastie, 1993).

One notable alternative theoretical account of courtroom decision making is the story model, developed by Nancy Pennington and Reid Hastie (1986, 1988; for reviews, see Pennington & Hastie, 1993, 1994). Rather than conceptualizing the decision-making process as a mathematical computation of numerically expressed evaluations of evidence, the story model focuses on the cognitive representation of the task. People actively construct narrative structures through which the evidence is organized, elaborated, and interpreted. Different stories correspond to different interpretations of the evidence, which lead, in turn, to different verdicts. For example, the case used most frequently by Pennington and Hastie concerns a first-degree murder charge. At one extreme, some jurors used the testimony to construct a story of self-defense, whereas at the other extreme, some jurors used the same testimony to construct a narrative of intentional, premeditated murder.

The acceptance of a story is determined by four certainty principles: coverage, coherence, uniqueness, and goodness of fit (Pennington & Hastie, 1986). Designed specifically to explain juror decision making in legal trials, this research has shown that ver-

dicts are determined by the constructed stories; the stories are not merely a post hoc justification of the verdicts (Pennington & Hastie, 1988). Although the story model undoubtedly fits the kind of task in which it was researched—a complex case that centered on questions of human intentionality and causality—its applicability to a broader range of tasks is less clear. The current research provides further support for the model's core insight regarding the holistic nature of decision-making processes and to extend the applicability of that insight to a broader range of decision tasks.

Coherence-Based Reasoning

Cognitive consistency theories have recently been revitalized by the advent of connectionist theories of cognition and, in particular, parallel constraint satisfaction mechanisms (Holyoak & Thagard, 1989; McClelland & Rumelhart, 1986; Read & Marcus-Newhall, 1993; Rumelhart & McClelland, 1986; Shultz & Lepper, 1996, 1998; E. R. Smith, 1996; Thagard, 1989). The intuition that underlies this approach is that the mind shuns incoherent representations and constructs coherent ones in their place. *Coherent* representations are ones in which elements that are positively related to one another tend to wax and wane together. In models of this type, complex tasks are represented by networks in which the decision variables are interconnected by excitatory and inhibitory links representing positive and negative relations among the variables. Constraint satisfaction models provide the processing mechanism by applying a relaxation algorithm that settles the network into a stable state in which the asymptotic activation levels of the units define a set of highly coherent variables. Bidirectional activations enable units that are mutually supportive via excitatory connections (i.e., those that "go together") to become highly active and to collectively inhibit their rivals, thus becoming increasingly coherent with the emerging decision. In this way, the representation of the task drifts toward either one of two points of equilibrium (we assume here only two alternative decisions). At each point of equilibrium the subset of the task variables that support the emerging decision are strongly endorsed, whereas the variables that support the rejected decision are suppressed or rejected; in effect, the representations of the alternatives become *spread apart*. The bidirectional influences between related units play a critical role in allowing the system to impose a coherent interpretation on the overall situation (Holyoak & Thagard, 1989; Read et al., 1997; Read & Miller, 1994).

As suggested by Read and Miller (1994; Read et al., 1997), parallel constraint satisfaction mechanisms offer a fitting solution to the limitations that stifled cognitive consistency theories. Connectionist representations are capable of capturing rich and large conceptual structures and relating them to the person's background knowledge. Constraint satisfaction algorithms provide a more realistic and nuanced means of resolving consistency than the crude mathematical rules used by consistency theorists. Read and Miller (1994) used constraint satisfaction simulations to model the restoration of balance and the reduction of cognitive dissonance, Shultz and Lepper (1996, 1998) simulated various findings of dissonance research, and Holyoak and Simon (1999) simulated coherence effects in an abstract reasoning task. Parallel constraint satisfaction mechanisms have also been observed in a number of experimental settings, including social reasoning (Read & Marcus-Newhall, 1993; Read, Snow, & Simon, 2003), analogical reasoning (Spell-

man & Holyoak, 1992), attitude change (Spellman, Ullman, & Holyoak, 1993), multiattribute decision making (Simon, Krawczyk, & Holyoak, *in press*) and abstract legal reasoning (Holyoak & Simon, 1999; Read et al., 2003; Simon, Pham, Le, & Holyoak, 2001).

The primary objective of the current research was to shed empirical light on the nature of making decisions that are based on multiple probabilistic judgments and inferences. Specifically, we tested whether the process can be understood as governed by the gestaltian principles of structural dynamic models. Such a finding could support an extension of the classic cognitive consistency theories to larger, semantically richer, and more complex tasks than previously conceived. This examination will also help determine whether the process conforms to the syntactic rules of meaning invariance and valuation-integration independence or whether it is better characterized as a coherence-driven construction of the cognitive representation, as predicted by parallel constraint satisfaction theories.

The second research objective derives from the broadly shared observation that factual inferences are closely linked to the decision maker's background knowledge. Representations of decision tasks cannot be constructed exclusively from the evidence explicitly designated for the task without reference to the fact finder's personal knowledge structures, analogies to experienced and hypothetical episodes, and more (Read, 1987; Schank & Abelson, 1995; Wyer & Radvansky, 1999). For example, an evaluation of the reliability of an eyewitness' identification of a suspect depends on one's background knowledge about people's ability to identify other people.

We tested whether the reverse might also be true: that the decision maker's background knowledge is affected by the emerging decision. Ziva Kunda (1990) and others have shown that when faced with preferred inferential goals, people tend to access only a subset of their background knowledge or to construct skewed representations of it in ways that serve to justify the desired conclusion. For example, Klein and Kunda (1992) demonstrated that people who expect to interact personally with a stereotyped person (a schizophrenic) tend to report more favorable attitudes and beliefs about members of the group. If, as we propose, the decision-making process is guided by connectionist models of parallel constraint satisfaction, it is theoretically possible that the bidirectionality reaches beyond the variables specifically identified in the task and also affects the content of the accessed background knowledge. In other words, the second objective was to test whether the coherence that forms between the emerging verdict and the evidence affects not only the evaluation of the evidence itself but whether it also affects the representations of background knowledge.

In all the following experiments, participants were asked to evaluate a number of pieces of evidence and background knowledge, first in isolated vignettes that shared no apparent relationship and then again in the context of a legal case in which they were asked to role-play a juror, arbitrator, or legal intern. In the latter phase, participants were presented with evidence that contained a great deal of factual ambiguity with respect to the identity of the perpetrator and were asked to determine whether the defendant was guilty or innocent of the crime. The decision hinged solely on the identity of the perpetrator. With some variations among the conditions, the case was designed so that it contained a number of

isolated pieces of circumstantial probabilistic evidence, none of which were plausibly related to other pieces of evidence. If the decision-making process of this task can be characterized by a parallel constraint satisfaction process involving bidirectional reasoning, we would expect the evaluations of evidence to shift toward lending stronger support to the chosen verdict.

We provide evidence of this bidirectional reasoning process in five different ways. First, across all four studies, we found that participants' evaluations of the facts and of the beliefs shifted from the pretest to the posttest to become more coherent with their final verdict. Second, participants reported high levels of confidence in their verdicts despite the ambiguity and complexity of the case. High levels of confidence are consistent with lopsided representations of the task in which the evidence supporting the chosen verdict is deemed to be strong and the contrary evidence is deemed weak. Third, in the first two studies we examined the impact of two different factors that significantly affect the likelihood of reaching a guilty verdict and found that the evaluations of the evidence maintained coherence with the verdicts, regardless of the distribution of the verdicts. Study 1 was designed to show that similar coherence shifts would be obtained even when we varied the likelihood of conviction—by manipulating the strength of the evidence. Study 2 was designed to show that even when attitudinal differences lead participants to different conviction rates (higher for individuals who support the death penalty), their evaluation of the evidence maintains coherence with the verdicts, regardless of the distribution of the verdicts. Study 3 provides a fourth kind of evidence for bidirectional reasoning. It was intended to show that when the preferred verdict changes, the evaluation of the facts change along with it. Specifically, we tracked a group of participants who first leaned toward one verdict but decided on the opposite verdict after the introduction of new evidence. We found that the evaluation of the nonmanipulated evidence closely tracked these shifts in verdict preferences across time. Study 3 also had a second purpose: It was designed to offer evidence that coherence shifts occur predecisionally and are not simply post hoc justifications of the verdict. Finally, Study 4 provides additional evidence for the bidirectionality of reasoning, specifically, for reasoning from the verdict to the evaluation of the evidence. We directly manipulated participants' verdicts by assigning them to the role of an intern who is helping a judge write an opinion for a case that has already been decided either in favor of or against the defendant. This assignment to side affected participants' evaluation of the evidence.

Study 1: Comparison of Manipulations of DNA Evidence

All of the following studies concern a case against an individual, Jason Wells, charged with stealing money from his employer's safe. In this first study, we manipulated whether available DNA evidence either strongly incriminated or exonerated Jason, a manipulation that was expected to lead to disparate conviction rates. First, we predicted that participants' evaluation of the evidence and the related beliefs would shift over the course of the study to become more consistent with their final verdict. Second, we predicted that despite the disparate conviction rates, the coherence shifts that accompanied the guilty and innocent verdicts would be quite similar in the two evidence conditions. This would indicate that even when the distribution of convictions was manipulated by

a strong piece of evidence, the evaluations of the nonmanipulated evidence would shift to maintain coherence with the corresponding verdicts.

Method

Design overview. Participants first responded to a set of pretest questions embedded in a series of vignettes that were apparently unrelated to each other. Unbeknownst to the participants, the questions actually concerned facts and beliefs that were contained in a legal case, which was later presented to them. In addition to these vignettes, participants were presented with additional decoy questions that were unrelated to the case.

After responding to the pretest questions, participants performed a distracting task consisting of a number of verbal analogies. Following this, they were presented with the criminal case against the defendant Jason Wells. The evidence included a DNA match that either incriminated or exonerated Jason as the perpetrator of the crime.

After reading over the case, participants were asked to render a verdict and then to fill out the posttest questions on the same facts and beliefs that they responded to in the pretest. Some of the questions pointed to Jason's guilt and some pointed to his innocence. Some of these questions were about the facts of the case, and others were about background knowledge and beliefs pertaining to those facts. Participants were divided into convicts and acquitters on the basis of the verdict they rendered. Thus, the overall design of the study was a 2 (DNA evidence: incriminating vs. exonerating; between subjects) \times 2 (acquitters vs. convicts; between subjects) \times 2 (pretest questions vs. posttest questions; within subjects) \times 2 (guilt vs. innocence questions; within subjects) \times 2 (fact questions vs. belief questions; within subjects) analysis of variance (ANOVA). Analyses were done separately for the fact and the belief questions.

Participants. A total of 154 people participated in this study via the Internet. Links placed on the Web site about.com invited U.S. citizens over the age of 18 to participate. As an incentive, it was explained that each participant would be entered into a lottery for \$200, with odds of winning at 1/200. The age of the sample ranged between 18 and 60 years, with a mean age of 34 years. Participants were 79.6% female and 20.4% male; 83.0% were White, 6.5% were African American, 4.0% were Latino, 2.0% were Asian, and 2.6% were Native American, with the remainder either listing themselves as "other" or failing to respond. As to education level, 33.8% had education up to high school, 54.5% had at least some college, and 11.6% had an advanced degree.

Note that all of the studies reported in this article were done on the Internet. However, we have previously done the same kinds of studies with very similar materials with students in the University of Southern California Psychology Department subject pool and obtained quite similar results (Simon, Snow, Brownstein, & Read, 2002).

Materials and procedure. After reading and accepting the online informed consent, participants were administered two sets of materials, separated by the distractor task of 20 analogies. A final demographic questionnaire was the last measure completed.

The first instrument (the pretest) asked participants to read a series of unrelated vignettes involving a variety of social situations. For example, one of the vignettes read as follows:

Wendy works as a computer programmer for a large insurance company. One evening, after most of the employees had left, she was walking by the office of the accounting department. She noticed a man rushing into the office and leaving a bouquet of flowers on the desk of Jessica Myers, one of the company's accountants. Jessica is a very shy person, and has experienced great difficulty in forging romantic relationships with men. The next day Wendy noticed that Jessica was distraught. Jessica told her that there was no note on the flowers and that she's eager to learn who had left them for her. When she learned that Wendy had seen the man, she became visibly excited. Jessica said

that she suspected that the person might be Dale Brown, who works for a travel agency located on the ground floor of the building. Wendy offered to go down to the travel agency to see if she could recognize Dale. Jessica waited nervously in the office. When Wendy returned, she told Jessica that she recognized Dale as the man she saw that evening. Jessica immediately asked her how certain she was. Wendy responded that she was completely certain that Dale was the man. She added that she had seen Dale around the building once or twice before.

Each vignette was followed by one or two related factual questions and 11-point response scales ranging from -5 (*strongly disagree*) to +5 (*strongly agree*). For example, following this vignette, participants were asked to rate their agreement with the following statement by clicking the appropriate radio button: "Wendy's identification makes it likely that it was Dale who left the flowers on Jessica's desk."

Some of the factual evaluations were followed by questions probing the participant's general belief on the issue at hand. For example, following the question about Wendy's identification, participants were asked to rate their agreement with the statement, "In general, when people identify someone whom they've seen once or twice before the identifications are accurate."

Participants responded to 12 questions of this type, 6 of which concerned facts discussed in the vignette and 6 of which concerned general beliefs related to the vignette. After filling out the first instrument, participants were then taken to a new Web page that presented an intervening distractor task, 20 verbal analogies from the Miller Analogies Test. On completion of this task, they were then taken to a new Web page with the second instrument.

The second instrument presented participants with a legal case involving a construction company ("Big Buildings") employee named Jason Wells, who was accused of stealing \$5,200 from the company's safe. The case was set up as an arbitration in a disciplinary proceeding brought by the employer against Jason. Participants were instructed to play the role of arbitrator whose task was to decide in favor of Jason Wells or the company. Participants then received some general information about Jason Wells followed by the evidence in the case and the arguments presented by the prosecution and defense.

The evidence included information about seven specific pieces of evidence, which were virtually unrelated to one another. One piece of evidence involved the presence of DNA material at the crime scene. In one version of the study, the DNA supported Jason Wells's guilt, whereas in the other version, the DNA supported Jason Wells's innocence. Of the remaining six pieces of evidence, three of the items tended to incriminate Jason: an eyewitness identification, a photo of a car like the one he drives leaving the parking lot around the time of the crime, and the possibility that he harbored a grudge against the company. Questions about these three items are labeled "guilt" questions. The other three items had an exonerating effect: He was seen far away from the crime scene later that evening, a possibility that he was working harder to make amends with the company, and a possibility that payments he made after the crime were legitimate family transactions. Questions about these pieces of evidence are labeled "innocence" questions.

A crucial part of the experimental design is that the evidentiary issues in the legal case, excluding the DNA evidence, were very similar to the vignettes evaluated in the pretest. For example, at trial, the following evidence was presented regarding the testimony of an eyewitness:

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

The DNA evidence was manipulated as follows:

Incriminating evidence: “The investigator testified that a tiny piece of skin was found on the handle of the safe. A DNA analysis of the skin specimen revealed that it matched Jason Wells’ DNA. The laboratory report stated that the odds were less than one to 7 million that the skin came from anyone other than Jason. The investigator testified that the skin specimen must have belonged to the person who stole the money because the safe had been thoroughly cleaned that evening, just minutes before the theft occurred.”

Exonerating evidence: “The investigator testified that a tiny piece of skin was found on the handle of the safe. A DNA analysis of the skin specimen revealed that it did not match Jason Wells’ DNA. The laboratory report stated that the odds were less than one to 7 million that the skin came from Jason. The investigator testified that the skin specimen must have belonged to the person who stole the money because the safe had been thoroughly cleaned that evening, just minutes before the theft occurred.”

After reading over the summary of the case and the evidence, participants read the company’s arguments and the defense’s arguments and were asked to decide on the verdict (a dichotomous choice between “convict” and “acquit”) and then rate their level of confidence in that verdict. Confidence ratings were made on an 11-point scale from 1 (*low confidence*) through 11 (*high confidence*). Participants were then asked to rate their evaluation of seven pieces of evidence that were presented at trial, using an 11-point scale ranging from -5 (*strongly disagree*) to +5 (*strongly agree*) (posttest). With the exception of a question assessing participants’ rating of the DNA evidence, the other questions were virtually identical to the ones asked about the vignettes in the pretest. For example, the question regarding the eyewitness testimony read as follows: “The technician’s identification of Jason makes it likely that the person hurrying out of the bookkeeper’s office was in fact Jason.”

As in the pretest, the factual determination of the eyewitness testimony was followed by the same question regarding the general belief about eyewitness testimony (e.g., “In general, when people identify someone whom they’ve seen once or twice before the identifications are accurate”). In total, participants responded to 12 questions that paralleled the questions asked during the pretest: 6 questions were about the actual facts of the case

(“fact” items), and 6 questions concerned the background beliefs (“belief” items). Within each category, 3 of the questions pertained to guilt items and 3 pertained to innocence items. The Appendix provides a complete list of the questions. Participants also responded to a question about the strength of the DNA evidence, as a manipulation check, but this question was not included in the analysis of the other items. Finally, participants were asked a number of demographic questions.

Results

Verdicts. As expected, manipulating the DNA evidence had a strong effect on verdicts. Of the participants who received the incriminating DNA evidence, 69.6% decided against Jason, whereas only 32.0% decided against him when they received the exonerating DNA evidence, $\chi^2(1, N = 154) = 21.80, p < .001$.

Confidence. We predicted that because of coherence with the chosen verdict, participants would display high levels of confidence in their decision, regardless of whether it was to convict or to acquit. High levels of confidence are an indicator of constraint satisfaction processing because they are the natural consequence of the spreading apart of the subsets of evidence, with the evidence supporting the chosen decision dominating the remainder of the evidence. Thus, despite the overall ambiguity of the case and the expected lack of complete agreement among participants on the verdict, we predicted that few, if any, participants would report low levels of confidence in their verdict. To make this easier to see, we combined participants’ verdicts and 11-point confidence scores to create a new confidence variable that ran from -11 (*highly confident of innocence*) to +11 (*highly confident of guilt*). As manifested by the highly bimodal distribution in Figure 1, our hypothesis was borne out. Less than 15% of the participants gave confidence ratings with absolute values of between 1 and 5, with the remainder all concentrated at levels of 6 to 11.

Coherence shifts. We next examine whether there is evidence for the expected bidirectional reasoning due to constraint satisfaction processes. The central hypothesis of the study was that the evaluation of the facts would shift to greater coherence with the eventual verdict. Thus, from pretest to posttest, convicts should give higher ratings on the guilt questions and lower ratings on the

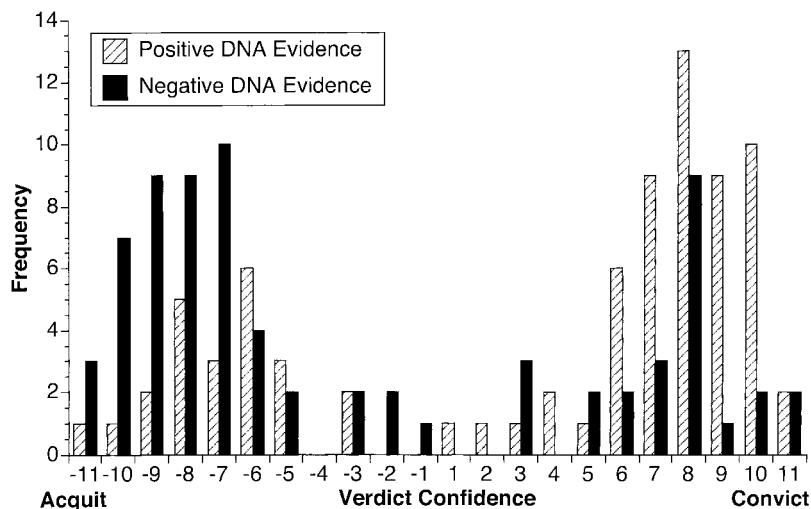


Figure 1. Study 1: Distribution of verdict confidences for both DNA conditions.

innocence questions. Conversely, acquiters should provide higher ratings on the innocence questions and lower ratings on the guilt questions. In contrast, classic models of evidence integration—IIT and Bayes's theorem—would predict no systematic shift in the ratings. We tested our prediction by analyzing the interaction among Pretest Versus Posttest Questions \times Guilt Facts Versus Innocence Facts \times Convictors Versus Acquitters. As predicted, this interaction was highly significant, $F(1, 150) = 66.30, p < .001$. Note that this analysis looked only at the fact items.

We then looked at the two-way interactions between pretest versus posttest and guilt facts versus innocence facts, separately for convictors and acquiters, to determine whether the opposing patterns of shifts for convictors and acquiters were significant. As predicted, both were significant. As seen on the left panels of Figure 2, for participants who decided against Jason ("Convictors"), the evaluation of the guilt questions increased relative to their own evaluations of virtually the same evidence on the pretest, and their evaluations of the innocence questions dropped dramatically, $F(1, 77) = 68.98, p < .001$. Opposite shifts are found for the participants who decided in Jason's favor ("Acquitters"), as seen in the right panels of Figure 2, $F(1, 73) = 10.84, p < .001$. In sum, the evaluations of the facts spread apart with the effect of one construction of the facts dominating the other, in line with the verdict. This domination, we believe, enables making the decision with confidence.

In addition to the hypothesized three-way interaction, there were also several significant main effects and interactions. There was a significant pre–post main effect, $F(1, 150) = 25.77, p < .001$, indicating a significant decline from pre to post in agreement with the fact items, as well as a significant guilt versus innocence items main effect, $F(1, 150) = 4.04, p = .046$, indicating overall greater agreement with the guilt items. Also, there were several significant

two-way interactions. One was between pre–post and verdict, $F(1, 150) = 68.81, p < .001$, indicating that convictors and acquiters showed different amounts of overall change from pretest to posttest, and another was between pre–post and guilt versus innocence questions, $F(1, 150) = 11.59, p < .001$, indicating that endorsement of guilt items was essentially unchanged from pre to post, whereas endorsement of innocence items declined from pre to post.

We hypothesized that a shift similar to that found on the facts would also occur in the background knowledge beliefs that are accessed in the course of a decision-making task. This hypothesis was confirmed, as indicated by the highly significant interaction among Pretest Versus Posttest \times Guilt Versus Innocence Beliefs \times Convictors Versus Acquitters, $F(1, 150) = 26.15, p < .001$. Figure 3 shows coherent shifts in the participants' general beliefs, which parallel the shifts in participants' evaluations of the evidence in the case. We also looked at the two-way interaction between pre–post and guilt versus innocence beliefs, separately for convictors and acquiters, to examine whether the two opposing patterns of shifts were significant. The two-way interaction for convictors was marginally significant, $F(1, 77) = 3.51, p = .065$, with little change in endorsement of guilt beliefs from pretest to posttest and a decline in endorsement of innocence beliefs. The acquiters displayed a significant opposing pattern, $F(1, 73) = 23.48, p < .001$, with a slight increase in endorsement of innocence beliefs and sharp decline in endorsement of guilt beliefs.

In addition to the hypothesized three-way interaction, there were also several significant main effects and interactions. There was a significant effect of verdict, such that convictors were generally more likely to endorse all belief statements than were acquiters, $F(1, 150) = 8.16, p = .005$. Second, there was a significant guilt versus innocence items main effect, $F(1, 150) = 13.62, p < .001$,

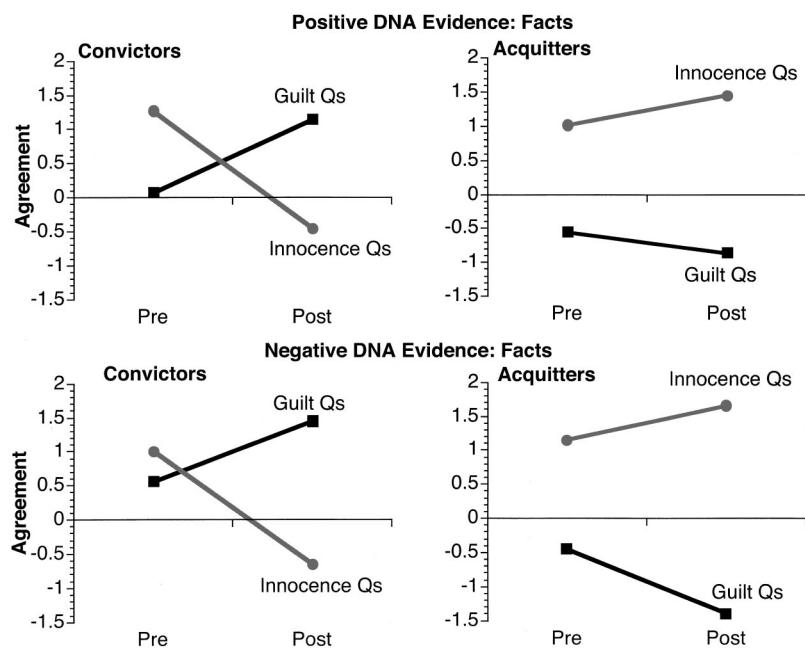


Figure 2. Study 1: Evaluation of guilt and innocence facts by acquiters and convictors for both DNA conditions. Qs = questions; Pre = pretest; Post = posttest.

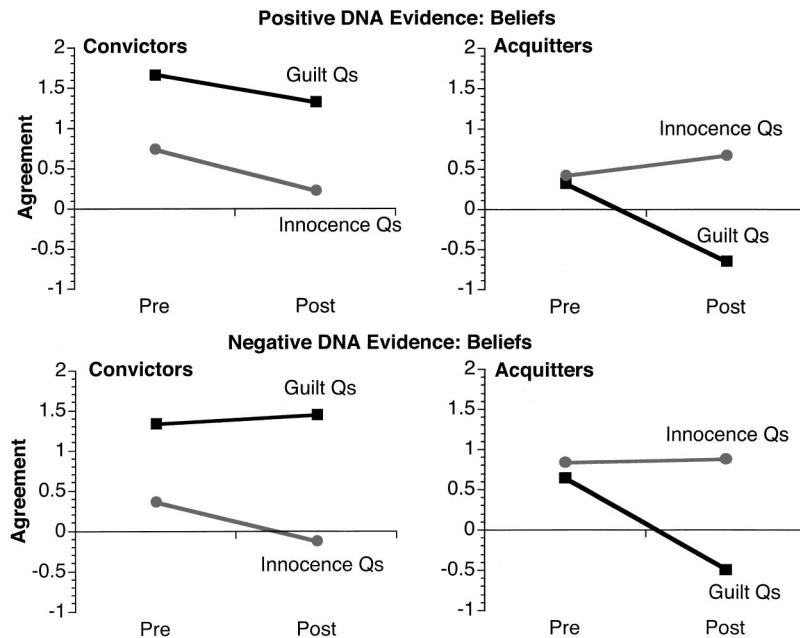


Figure 3. Study 1: Evaluation of guilt and innocence beliefs by acquittees and convictors for both DNA conditions. Qs = questions; Pre = pretest; Post = posttest.

indicating that there was overall greater agreement with the guilt items. Third, there were several significant two-way interactions. One was between pre-post and the guilt versus innocence items, $F(1, 150) = 6.95, p = .009$, because endorsement of guilt items declined more sharply from pre to post than did endorsement of innocence items. There was also a significant interaction between pre-post and verdict, $F(1, 150) = 35.37, p < .001$, with acquittees showing greater decline in endorsement of the belief items from pretest to posttest.

We hypothesized that essentially identical coherence shifts would occur, irrespective of changes in conviction rate caused by the different strengths of evidence. As noted above, changing the direction of the DNA evidence had a strong effect on the distribution of verdicts. However, this change in verdict distributions seemed to have had no effect on the representations of the case by the individuals in the two different verdict groups, as can be seen by an examination of the impact of the DNA evidence on the ratings of the nonmanipulated facts and the beliefs. As can be seen in Figures 2 and 3, the evaluations of the evidence within each group of verdicts were not affected by the strength of the case. The patterns of evaluation of both the facts and the beliefs are quite similar for the two different DNA conditions, despite the disparate distribution of convictions and acquittals. This is corroborated by the complete lack of any interaction of the DNA evidence manipulation with the coherence shift effect (four-way interaction among pre-post, guilt vs. innocence questions, verdict, and strength of case); fact questions, $F(1, 150) = 0.26, p = .612$; and belief questions, $F(1, 150) = 0.38, p = .538$. Once participants reached a verdict, they ended up with a pattern of coherence that is typical for that verdict.

Coherence shifts and confidence. We also expected that participants' confidence in their verdicts would be related to the size

of their coherence shifts. We examined this in two ways. First, we found that the greater the absolute value of the difference between endorsement of the guilt and innocence items at posttest, the more confident participants were in their verdicts, $r(151) = .395, p < .001$. Second, we also expected that confidence should be related to the size of the shift from pretest to posttest. To test this, we first calculated an overall coherence shift score, creating what we called a guilt composite by reverse coding the innocence items so that greater endorsement was consistent with guilt and then adding the innocence and guilt items together. We then calculated the absolute value of the change in the guilt score from pretest to posttest and found that participants with stronger coherence shifts were also more confident, $r(149) = .329, p < .001$.

Thus, as predicted, we found strong evidence for coherence-based reasoning in arriving at a verdict. Participants showed substantial shifts from pretest to posttest in their evaluations of both the facts of the case and their background beliefs, and they reported high levels of confidence. Moreover, the form of these coherence shifts was not affected by shifts in the distribution of verdicts. Overall, these results contradict the two principles of meaning invariance and value-integration independence that are fundamental assumptions of IIT and Bayesian models. These results also provide some support for our contention that the reasoning process proceeds bidirectionally, specifically, that it has a "backward" component by which decisions affect the perception of the evidence. It is important to keep in mind that it is inconceivable that the manipulated evidence could bear any direct influence on the other pieces of evidence, because a DNA match has no plausible relation to the accuracy of the eyewitness' identification or to the likelihood that the defendant could have made it in time to the place where he was seen later that evening. It is even less conceivable that the specific DNA match could have any

influence on one's beliefs about witnesses' ability to identify people in general or on whether one can save much time by driving aggressively through Los Angeles evening traffic. There is good reason to believe that the influence of the manipulated evidence on the nonmanipulated facts and beliefs was mediated through the verdicts: The DNA manipulation influenced participants' verdicts, and constraint satisfaction processing drew the evidence toward coherence with these verdicts. The impact of emerging verdicts on the evaluation of the evidence was tested more directly in Study 4.

Study 2: Comparing Participants With Different Attitudes

In this study, we manipulated whether Jason was charged only with theft or whether he was charged with murder for killing a security guard during the commission of the theft. We expected that participants asked to decide on a murder charge would have a lower conviction rate because of the greater penalty for conviction (e.g., Kerr, 1978). In addition, we also measured participants' attitudes toward the death penalty. Previous research (e.g., Cowan, Thompson, & Ellsworth 1984; Goodman-Delahunty, Greene, & Hsiao, 1998) has demonstrated that individuals with stronger pro-death-penalty attitudes are more likely to convict defendants in criminal trials. We expected that both the severity of the crime and death penalty attitudes would influence the conviction rate. However, as in the previous study, we expected that evaluations of the evidence would shift in line with the verdicts, with the result that the changes in conviction rate would not influence the patterns of coherence shifts.

Method

Design overview. The same basic procedure and materials as in Study 1 were used in this study with the following changes. First, after completing the analogy distractor task, participants were presented with a scale on attitudes toward the death penalty (O'Neil, Patry, & Penrod, in press), which was used to divide participants into three groups of strong pro-death-penalty attitudes, strong anti-death-penalty attitudes, and moderate attitudes on the death penalty. Second, half of the participants were presented with the theft version of the Jason materials, similar to the one used in Study 1, with the difference that no DNA evidence was provided; instead, evidence was provided about a call made on Jason's cell phone from near the scene of the crime just minutes after the money was stolen. The other half of the participants were presented with a version in which a guard is killed during the theft and as a result, Jason is being charged with murder. Both charges were based on the identical set of evidence, all of which was circumstantial. Third, the case was set up as a criminal trial, and the participants were asked to role-play a juror.

As in the previous study, after reading over the case, participants were asked to render a verdict and then to fill out posttest questions on the same facts and beliefs that they responded to in the pretest. Participants were divided into convicts and acquittees on the basis of the verdict they rendered. Thus, the overall design of the study was a 2 (type of case: theft vs. murder; between subjects) \times 3 (death penalty attitudes: anti-death penalty, moderate, pro-death penalty; between subjects) \times 2 (acquittees vs. convicts; between subjects) \times 2 (pretest questions vs. posttest questions; within subjects) \times 2 (guilt vs. innocence questions; within subjects) \times 2 (fact questions vs. belief questions; within subjects). However, analyses were done separately for the fact and the belief questions.

Participants. Three hundred forty-four people participated in this study via the Internet.¹ They were recruited as in Study 1. The ages of the participants ranged from 18 to 68 years, with a mean of 36 years; 78% of

the sample was female, and 22% was male. We did not gather further demographic information for this sample.

Materials and procedure. The materials used in this study were similar to those used in the earlier study, with the following differences. First, when participants filled out the distractor task of 10 analogies, they also filled out O'Neil et al.'s (in press) Attitudes Toward the Death Penalty scale.

Second, in contrast to the previous study, in which the case against Jason was portrayed as an arbitration in a disciplinary proceeding brought by the employer against Jason, here the case was portrayed as a criminal case against Jason. After doing the distractor task but before reading any information about the case, participants were instructed to play the role of a juror in a criminal case whose task was to decide whether or not Jason was guilty of the crime with which he was charged. Participants were then immediately given the beyond a reasonable doubt standard of proof to use when considering their verdict. After reading the definition of the standard of proof, participants were asked to indicate their interpretation of the standard by clicking a radio button that corresponded (in increments of 5) to a number from 1 to 100 to indicate the level of confidence they would need to find Jason guilty on the basis of the prescribed standard of proof.

Third, participants received either a theft version or a murder version of the case, which were essentially identical with the exception that in the murder case it was stated that a guard had been brutally beaten during the commission of the theft and had died from the beating, resulting in a murder charge against Jason. Before making the verdict decisions, participants were again given the beyond a reasonable doubt standard of proof. They were then asked to reach their verdict, rate their confidence in their verdict, and then fill out the posttest questions to evaluate the evidence. Following the posttest, participants were again given the standard of proof and asked to provide their interpretation of it by indicating the level of confidence needed to find Jason guilty on the basis of the prescribed standard of proof. Finally, participants were asked a number of demographic questions. No significant effects were found for the interpretations of standards of proof or demographic data, so these data are not discussed further.

Results

This study examined whether attitudes toward the death penalty and type of criminal case (theft vs. murder) affected the ways in which acquittees and convicts represent the evidence and beliefs in the case. Because the type of criminal case did not significantly affect the conviction rate, $\chi^2(1, N = 344) = 2.91, p = .088$, we focus here on the impact of death penalty attitudes, collapsing over type of criminal case. In examining the impact of death penalty attitudes, we present further evidence for both constraint satisfaction processes and the bidirectionality of reasoning.

Verdicts. Death penalty attitudes were related to conviction rates: 21.9% of the anti-death-penalty individuals found against Jason, 34.7% of the moderates found against Jason, and 43.9% of the pro-death-penalty individuals found against Jason, $\chi^2(2, N = 344) = 10.616, p = .005$. Note that the pro-death-penalty individuals were twice as likely to convict as were the anti-death-penalty individuals. As in Study 1, participants were highly confident in

¹ We took several steps to ensure that subjects only participated in one of our studies. First, in the instructions to all our studies, we clearly told subjects that they could not do the study if they had done it before and that only new participants would be eligible for the lottery. Second, we collected both the e-mails and the Internet protocol addresses of the data submissions, and from examining them it is clear that less than 1% of our sample might have been repeats.

their decisions. Less than 10% of the participants gave confidence ratings of between 1 and 5, with the remainder all concentrated at levels of 6 to 11.

Coherence shifts. As in Study 1, we hypothesized that the evaluation of the facts would shift to greater coherence with the eventual verdict, with convictors providing higher ratings on the guilt questions and lower ratings on the innocence questions, and vice versa for the acquittees. This was tested by the interaction among Pretest Versus Posttest Questions \times Guilt Facts Versus Innocence Facts \times Convictors Versus Acquittees. As predicted, this interaction was highly significant, $F(1, 255) = 85.25, p < .001$.

As in the previous study, we examined the two-way interactions between pre–post and guilt versus innocence items, separately for convictors and acquittees. Both were highly significant. As seen in the left panels of Figure 4, convictors' evaluation of the guilt questions increased slightly relative to their own evaluations of virtually the same evidence on the pretest, and their evaluations of the innocence questions dropped dramatically, $F(1, 88) = 60.12, p < .001$. In contrast, acquittees showed a sharp decline for guilt items and were essentially unchanged for the innocent items, as seen in the right panels of Figure 4, $F(1, 173) = 47.14, p < .001$. In sum, the evaluations of the facts spread apart, with one construction of the facts dominating the other.

However, this pattern of coherence shifts was unaffected by the participants' death penalty attitudes, as indicated by the complete lack of any interaction of the death penalty attitudes with the coherence shift (four-way interaction among pre–post, guilt vs. innocence questions, verdict, and death penalty attitudes), $F(2, 255) = 1.66, p = .192$. As can be seen in Figure 4, the evaluations of the evidence within each group of verdicts are not affected by death penalty attitudes. The patterns of evaluation of the facts are almost identical for the anti- and pro-death-penalty participants,

despite the substantially disparate distribution of convictions and acquittals. Once again, it appears that when participants reach a verdict, they achieve a pattern of coherence that is characteristic for that verdict.

In addition to the significant hypothesized three-way interaction, there were several significant main effects and interactions. First, there was a significant pre–post main effect, $F(1, 255) = 75.20, p < .001$, indicating a significant decline from pre to post in agreement with all the fact items. Second, there was a significant guilt versus innocence items main effect, $F(1, 255) = 10.87, p = .001$, indicating that there was overall greater agreement with the guilt items. Third, there were several significant two-way interactions. One was between guilt versus innocence items and death penalty attitudes, indicating that as death penalty attitudes increased, participants were more likely to endorse guilt items and less likely to endorse innocence items. The significant interaction between guilt versus innocence items and verdict, $F(1, 255) = 86.94, p < .001$, was due to a greater endorsement of guilt items by convictors and a greater endorsement of innocence items by acquittees.

As in Study 1, a similar coherence shift occurred in the background knowledge beliefs. However, in the interests of space, we do not discuss these findings further.

Coherence shifts and confidence. Again, participants' verdict confidence was related to the size of their coherence shifts. First, the greater the absolute value of the difference between endorsement of the guilt and innocence items at posttest, the more confident participants were in their verdicts, $r(276) = .315, p < .001$. Second, participants with stronger coherence shifts, as measured by the absolute value of the change in the guilt composite from pretest to posttest, were also more confident, $r(271) = .172, p = .004$.

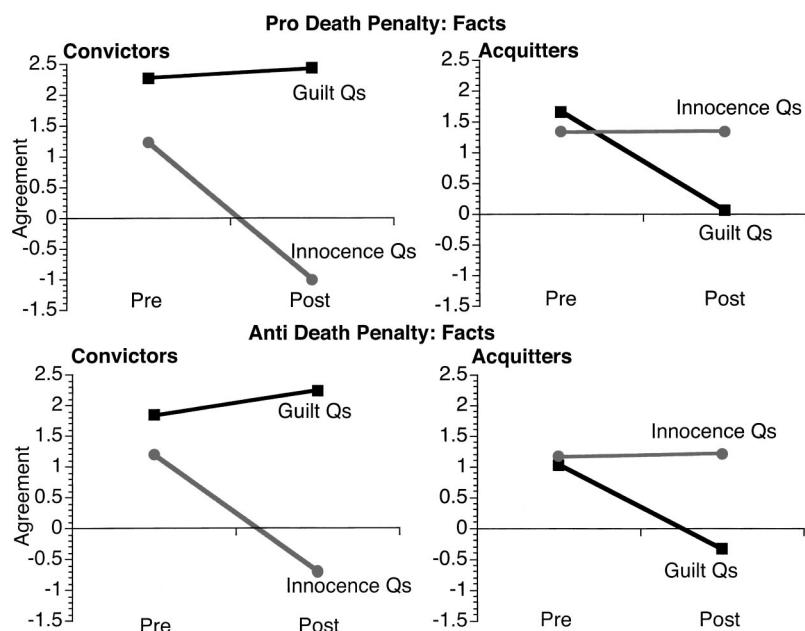


Figure 4. Study 2: Evaluation of guilt and innocence facts by acquittees and convictors with pro- or anti-death-penalty attitudes. Qs = questions; Pre = pretest; Post = posttest.

Again, we find strong evidence of coherence-based reasoning, because participants' evaluations of the facts and beliefs shifted to become coherent with their verdict. And again, a factor that affects the distribution of verdicts, death penalty attitudes, had no effect on the form of the coherence shifts.

Study 3: Effects of the Introduction of New Evidence and Predecisional Measurements

This study had two goals. First, we wanted to provide a somewhat different form of evidence for bidirectionality in reasoning. Specifically, we hoped to show that participants who shifted their preference from one verdict to the other would also reverse their evaluation of the evidence. Second, we hoped to provide additional evidence to support the view that coherence shifts happen predecisionally rather than as post hoc justifications for already made decisions (see Simon et al., 2001, *in press*). That is, we wanted to establish that even before committing themselves to a verdict, participants' evaluation of the evidence begins to shift toward their eventual verdict.

The basic design of this study is similar to the previous studies. However, testing the above hypotheses necessitated some changes in procedure. Most importantly, we introduced an intervening measurement session between the pretest and posttest phases. As in the previous studies, participants first read over the various scenarios and responded to the pretest questions and then completed the intervening task. Before being presented with the case of Jason Wells, participants were told that an important piece of evidence was not yet available and was expected to be obtained at a later stage. They were asked to read the case but were advised to withhold making any decision on the verdict until the missing evidence became known. After reading over the case, participants were reminded that they should continue to suspend their decision until the missing evidence was made available, and they were assured that any assessments they would make at this point would not be binding. Specifically, the instructions read as follows:

It is important that you remember that there is no point in making a decision before obtaining the missing piece of evidence. Any assessments that you make at this point remain private, and they do not bind you in any way. Later on, after having received the evidence in its entirety, you will be free to make whichever decision you think is appropriate.

Participants were advised to think of the case in the meantime; they were then asked to indicate their initial feelings or "leanings" toward either one of the verdicts, and then were asked to indicate their agreement with the facts and beliefs associated with the case.

Immediately after indicating their leanings and making the other ratings, participants were taken to a new Web page that presented them with one of three possible types of new information. In one condition, they were given a new piece of evidence that favored guilt ("conviction information"): Testimony was presented that Jason had made a 4-min call on his cell phone from several blocks away from his office only several minutes after the crime had occurred. In a second condition, they were given a new piece of evidence that favored Jason's innocence ("acquittal information"): Testimony was presented that Jason had made a 4-min call on his cell phone shortly after the crime was committed from a location near his house, which was about 45 min from the scene of the

crime. In the third condition, they were told that the expected new evidence was not obtained and that they should decide the case with the available evidence. After reading the new evidence (or being informed that no new evidence was available), participants were notified that the evidence was complete. Prior to eliciting participants' verdicts and evaluations, the instructions read, "You are free to decide either way, regardless of any tentative leanings you might have had earlier on." Participants were then asked to indicate their final verdict and to evaluate the same facts and beliefs that they had evaluated during the intervening session.

Method

Design overview. Participants received one of three types of new information (acquittal, conviction, or no new information) and gave their evaluations of the evidence and the beliefs at three different times: pretest; after initially reviewing the evidence but before making a decision and while awaiting the missing evidence (leaning); and at posttest, after receiving the new information and rendering a final verdict. Thus, the overall design of the study was a 3 (type of new information: acquittal, conviction, or no new information; between subjects) \times 3 (time of measurement: pretest, leaning, posttest; within subjects) \times 2 (guilt vs. innocence questions; within subjects) \times 2 (fact questions vs. belief questions; within subjects) factorial. As in the previous studies, the fact and belief questions were analyzed separately.

Participants. Two hundred eighty-six people participated in this study via the Internet. They were recruited as in the previous studies. Ages ranged from 18 to 74 years ($M = 36$), and 82% of the sample was female.

Results

We measured participants' verdict tendencies after they first received the case but before they had committed themselves to a verdict (leanings). Further, after participants provided their initial leanings but before they reached a final verdict, we manipulated whether they received new information that was consistent with conviction, consistent with acquittal, or did not receive new information. Following this, we measured their final verdict. Because we had measured initial leanings, we could determine whether participants changed their mind from their initial leanings ("switchers") or maintained their initial leanings. In the following analyses, we were interested in tracking the evidence evaluations for those who maintained a consistent position versus those who switched their positions.

Confidence. As in the previous two studies, once participants arrived at a verdict they were typically highly confident in their decision. Figure 5 provides the verdict confidence scores both for individuals who initially leaned toward conviction and those who initially leaned toward acquittal. Clearly, even individuals who started out leaning in one direction but then switched to a different verdict (in Figure 5, "Lean Acquit" on the "Convict" side of the distribution and "Lean Convict" on the "Acquit" side of the distribution) tended to be confident in their final verdict. Although one might expect that individuals who switched between their tentative verdict and their final verdict would tend to be uncertain in their decision, that was typically not the case.

Coherence shifts. We present two sets of analyses of the evidence evaluations. The first set focuses on participants who began with a leaning toward conviction at the middle phase and compares those who consistently maintained a conviction verdict

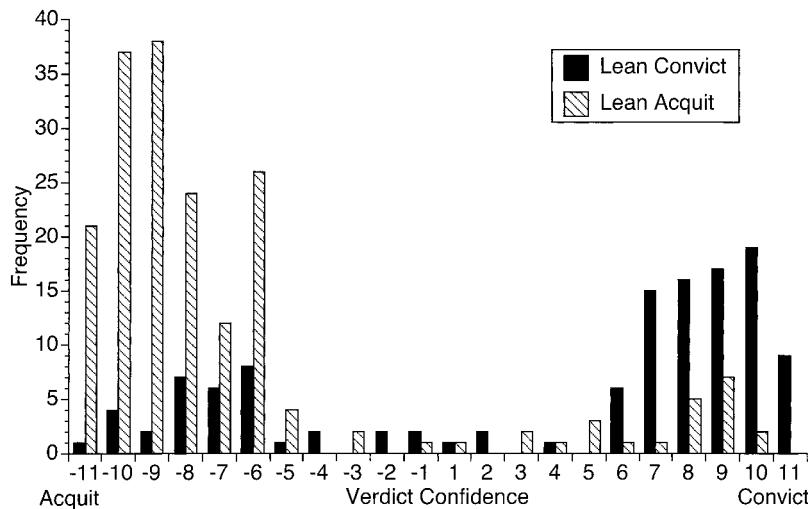


Figure 5. Study 3: Distribution of verdict confidences for individuals who initially leaned toward conviction or acquittal.

("consistent convictors") with those who ultimately switched to acquittal ("acquittal switchers"). The second set of analyses focuses on participants who began with a leaning toward acquittal at the middle phase and compares those who consistently maintained an acquittal verdict ("consistent acquitters") with those who ultimately switched to guilt ("guilt switchers"). Note that these analyses are focused comparisons of subsets of the entire design.

Consistent convictors versus acquittal switchers. Of the 112 participants in the first set of analyses, 79 maintained a consistent preference for conviction, whereas 33 initially leaned toward conviction but then switched to acquittal. Although the majority of the acquittal switchers were participants who had received acquittal information ($n = 17$), a small number had received either no new information ($n = 10$) or conviction information ($n = 6$). We decided to include all the switchers in this group, regardless of why they switched, because our main focus was on how evidence was interpreted once a switch occurred. The analyses were performed as a 2 (consistent convictors vs. acquittal switchers; between subjects) \times 3 (type of new information presented: conviction information, acquittal information, no new evidence; between subjects) \times 3 (time of measurement: pretest, leanings, posttest; within subjects) \times 2 (type of questions: guilt vs. innocence; within subjects) mixed ANOVA.

As predicted, for the evaluations of the facts, we obtained the three-way interaction among time of measurement, guilt versus innocence questions, and consistent convictors versus acquittal switchers, $F(2, 105) = 13.33, p < .001$. As can be seen in the left side of Figure 6, the consistent convictors showed an increasing spreading over time in their evaluation of the guilt and innocence items, giving higher evaluations to the guilt questions and lower evaluations to the innocence items. This pattern resembles the shifts we observed in the two previous studies. In contrast, the pattern of the acquittal switchers is different. Their evaluations first shifted from pretest levels to a state of coherence with their initial leaning of guilt (higher evaluation for the guilt items and lower evaluation for the innocence items), but then shifted in the opposite direction toward coherence with their final verdict of

innocence (a higher evaluation for the innocence items and a lower evaluation for the guilt items). This pattern indicates that the evaluation of the evidence dovetails with the preferred verdict. Note also that both the consistent convictors and the acquittal switchers showed spreading of the evidence at the leaning measurement. Thus, coherence shifts happened well before a commitment was made to a decision.

In addition to the predicted three-way interaction, we also obtained several other main effects and interactions. There was a significant main effect for time of measurement, $F(2, 105) = 5.90, p = .004$, indicating that the average evaluations changed over time. In addition, participants gave greater endorsement to the guilt questions compared with the innocence questions, $F(1, 106) = 8.41, p = .005$.

There was also a significant interaction of guilt versus innocence questions with the type of information participants received, $F(2, 106) = 5.84, p = .004$, such that the difference in endorsement of guilt and innocence questions was smaller for participants who received convict information compared with participants who received the other two types of information. Further, there was a significant interaction of type of questions with the final verdict, $F(1, 106) = 43.22, p < .001$, indicating that consistent convictors more strongly endorsed the guilt items and the acquittal switchers more strongly endorsed the innocence items. In addition, the type of questions also interacted with the time of measurement, because the difference between endorsement of guilt and innocence questions increased over time, $F(2, 105) = 16.58, p < .001$. Finally, there was a three-way interaction between type of questions, time of measurement, and the type of new information presented, $F(4, 212) = 4.91, p = .001$, that was difficult to interpret.

As in the previous studies, similar effects were found on the belief items. The consistent convictors showed a spread from pretest to posttest between the guilt and innocence beliefs in a direction consistent with their verdict. Acquittal switchers did not show such a shift. However, to conserve space we do not discuss these results further.

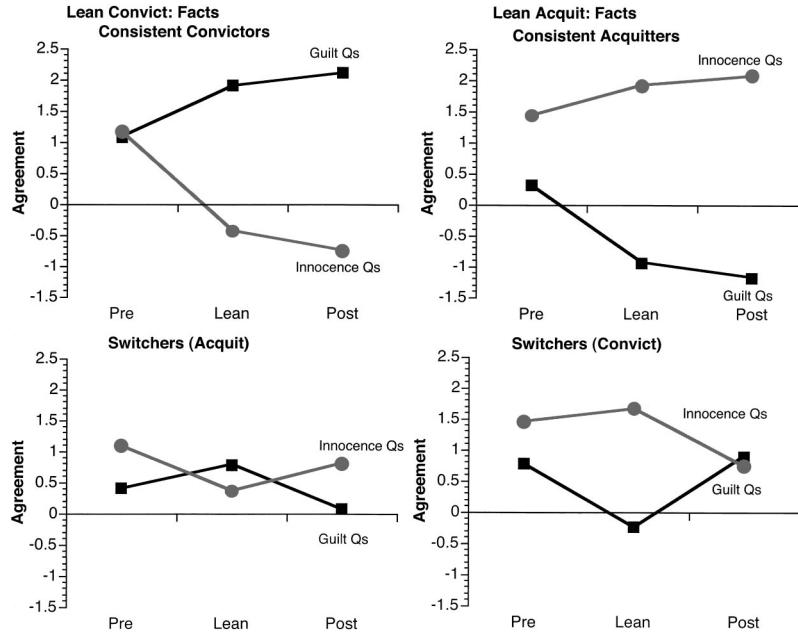


Figure 6. Study 3: Evaluation of guilt and innocence facts by consistent convictors compared with switchers from conviction to acquittal (left side) and by consistent acquitters compared with switchers from acquittal to conviction (right side). Qs = questions; Pre = pretest; Lean = preliminary leaning; Post = posttest.

Consistent acquitters versus guilt switchers. The second set of analyses focuses on participants who began with a leaning toward acquittal at the middle phase and compares those who consistently maintained an acquittal verdict (consistent acquitters) with those who ultimately switched to guilt (guilt switchers). Note that these analyses are focused comparisons of subsets of the entire design. One hundred fifty-two participants were consistent acquitters, whereas 22 initially leaned toward acquittal but then switched to conviction.

Although most of the guilt switchers were participants who had received guilt information ($n = 16$), a small number had received either no new information ($n = 2$) or acquittal information ($n = 4$). Again, we included all the guilt switchers in the same group, regardless of their reason for switching, because we were primarily interested in how the evidence was interpreted once a switch occurred. The analyses were performed as a 2 (consistent acquitters vs. guilt switchers; between subjects) \times 3 (type of new information presented: conviction information, acquittal information, no new evidence; between subjects) \times 3 (time of measurement: pretest, leanings, posttest; within subjects) \times 2 (type of questions: guilt vs. innocence; within subjects) mixed ANOVA.

As predicted, for the evaluations of the facts, we obtained the three-way interaction among time of measurement, guilt versus innocence questions, and consistent acquitters versus guilt switchers, $F(2, 167) = 7.48, p = .001$. As can be seen in the right side of Figure 6, the consistent acquitters showed an increasing spreading over time in their evaluation of the guilt and innocence items, giving higher evaluations to the innocence evidence and lower evaluations to the guilty evidence. This is the pattern found in the two previous studies. In contrast, the guilt switchers showed a pattern of first changing their evaluations of the evidence to be consistent with their initial leaning of innocence, with higher

evaluations for innocence items and lower for guilt items, but then switching by the final verdict to show equivalent evaluations for the guilt and innocence evidence. Again we see that the evaluation of the evidence dovetails with the participants' verdicts.

In addition to the predicted three-way interaction, we also obtained several other main effects and interactions. There was a significant main effect for time of measurement, $F(2, 167) = 3.45, p = .034$, indicating that the average evaluations changed over time. In addition, participants gave greater endorsement to the innocence evidence compared with the guilt evidence, $F(1, 106) = 8.41, p = .005$.

There was also a significant interaction of type of questions with the final verdict, $F(1, 168) = 8.62, p = .004$, indicating that consistent acquitters more strongly endorsed the innocence evidence than the guilt evidence, whereas guilt switchers showed a smaller difference in the endorsement of innocence and guilt evidence. In addition, the type of questions also interacted with the time of measurement, because the difference between endorsement of guilt and innocence questions increased over time, $F(2, 167) = 4.24, p = .016$.

As in the previous studies, consistent acquitters exhibited a consistent shift in belief evaluations from pretest to posttest in a direction consistent with their verdict. Guilt switchers did not show such a shift. However, to conserve space we do not discuss these results further.

Both the acquittal switchers and the guilt switchers demonstrated that participants' evaluation of the evidence first shifted to cohere with their tentative verdicts and then shifted in the opposite direction toward coherence with their final verdicts. This shifting from the leaning measurement to the final measurement provides further evidence for the bidirectionality of reasoning. It should also be noted that as predicted, shifts in evidence evaluation occurred

before participants committed themselves to the final verdicts. Participants showed quite strong coherence shifts at the leaning measurement, before they had made their final decision and at a time when they were told that they should wait for possibly compelling new evidence and were urged to withhold their decisions. Thus, it appears that coherence shifts occur as part of the process of moving toward a verdict and not merely as a post hoc justification of it.

Study 4: Effect of Assignment to a Side on Coherence Shifts

The results of Studies 1–3 all support the postulated bidirectionality in reasoning between verdict and evidence evaluation. However, these studies relied on participants' choice of a verdict and thus were correlational. The case for bidirectionality would be strengthened if we could somehow randomly assign participants to a verdict of guilt or innocence and show that this affected the evaluation of the evidence. Here we addressed this issue by asking participants to role-play a legal intern whose job is to help a judge draft the supporting arguments for a verdict at which the judge has already arrived. Thus, some participants were assigned to help write an opinion supporting Jason Wells's guilt, and others were assigned to help write an opinion supporting his innocence.

Method

Participants. A total of 163 people participated via the Internet, with recruitment procedures identical to those in the previous studies. Participants ranged in age between 18 and 81 years ($M = 33$). The sample was 73.0% female and 74.1% White, 4.0% Latino, 5.7% African American, 4.0% Asian, and 1.1% Native American; the remainder described themselves as "other" or did not respond. In terms of education, 35.2% had up to a high school education, 48.9% had at least some college, 8.0% had an advanced degree, and 6.9% failed to answer.

Materials and procedure. The overall design and materials of this study were essentially the same as in Study 1. The only difference was that instead of including DNA evidence as a between-subjects manipulation, participants were asked to assume one of two possible roles as a legal intern. More specifically, participants were asked to play the role of a 3rd-year law student who was an intern for a retired judge who was serving as an arbitrator in labor disputes. The instructions were designed to encourage the participants to accept the correctness of the verdict to which they were assigned. The judge was described as highly respected, thorough, and fair. It was explained that the judge had already heard the evidence prior to the participants' commencement of the internship and that on the basis of the evidence presented, she had already decided either against Jason Wells or in his favor. Participants were told that the judge had assigned them to assist in writing the first draft of the opinion and had instructed them to read the transcript of the evidence and to think of reasons that justified the verdict. Participants were never actually asked to write this opinion, but they were told this in the attempt to get them to arrive at a verdict consistent with the judge's decision. After reading the case, participants were asked to give their own verdicts as well as respond to the belief and fact questions. We expected that the participant's verdict would be consistent with the judge's verdict. All other materials and questions were identical to Study 1.

Results

Before performing mediational analyses to examine whether assignment to a side results in a significant coherence effect,

mediated at least partially by participants' verdicts, we first briefly examined whether we obtained the same coherence shifts for facts and beliefs as in the previous studies and also whether assignment to a side resulted in a significant coherence shift.

Verdicts and coherence shifts. As in the previous studies, we found significant coherence shifts for the facts. The evaluation of the facts shifted to greater coherence with the eventual verdict, with convicts providing higher ratings on the guilt questions and lower ratings on the innocence questions, and vice versa for the acquittees (see Figure 7). This was tested by the interaction among Pretest Versus Posttest Questions \times Guilt Facts Versus Innocence Facts \times Convictors Versus Acquittees. As predicted, this interaction was highly significant, $F(1, 159) = 31.33, p < .001$. In the interest of conserving space, other effects are not reported.

As in the previous studies, a similar coherence shift occurred in the background knowledge beliefs (see Figure 8), as indicated by a significant three-way interaction: Pretest Versus Posttest Questions \times Guilt Beliefs Versus Innocence Beliefs \times Convictors Versus Acquittees, $F(1, 159) = 7.46, p = .007$. Here the coherence shifts seemed to occur largely for the acquittees. In the interest of conserving space, other effects are not reported.

Assignment to a side and coherence shifts. We also found significant effects of the assignment manipulation on coherence shifts for both the facts and beliefs, although they were somewhat weaker than the verdict effects. First, to test whether assignment to a side led to significant coherence shifts for the facts, a three-way ANOVA (without verdict as a factor) was performed (Pretest vs. Posttest Questions \times Guilt Facts vs. Innocence Facts \times Assignment to Big Buildings vs. Assignment to Jason Wells). This was significant, $F(1, 161) = 7.84, p = .006$. As can be seen in Figure 7, from pretest to posttest, convicts more strongly endorsed guilt facts and less strongly endorsed innocence facts. As expected, acquittees' endorsement of guilt facts decreased, although there was also a decrease in their endorsement of innocence facts.

Second, to test whether assignment to a side had a similar impact on coherence shifts for beliefs, a three-way ANOVA (without verdict as a factor) was performed on the belief questions (Pretest vs. Posttest Questions \times Guilt Beliefs vs. Innocence Beliefs \times Assignment to Big Buildings vs. Assignment to Jason Wells). This was significant, $F(1, 161) = 17.00, p < .001$. As can be seen in Figure 8, from pretest to posttest, acquittees more strongly endorsed innocence beliefs and less strongly endorsed guilt beliefs. Convicts showed slight declines in endorsement of both guilt and innocence beliefs.

Mediational analyses. We now examine whether the impact of assignment to a side was at least partially mediated by participants' verdicts. To do this, we performed a series of mediational analyses, using linear regression.

As predicted, being assigned to a particular side of the case did significantly influence participants' verdicts in the case. Of those assigned to help the judge write an opinion justifying her decision to convict Jason, 61.0% thought Jason was guilty, whereas only 39.5% of those assigned to help write an opinion justifying Jason's innocence thought Jason was guilty, $\chi^2(1, N = 163) = 7.52, p = .006$. Although assignment to a side did not completely determine participants' verdicts, it did have a substantial impact, and one strong enough to allow us to demonstrate that verdicts did influence the evaluation of evidence.

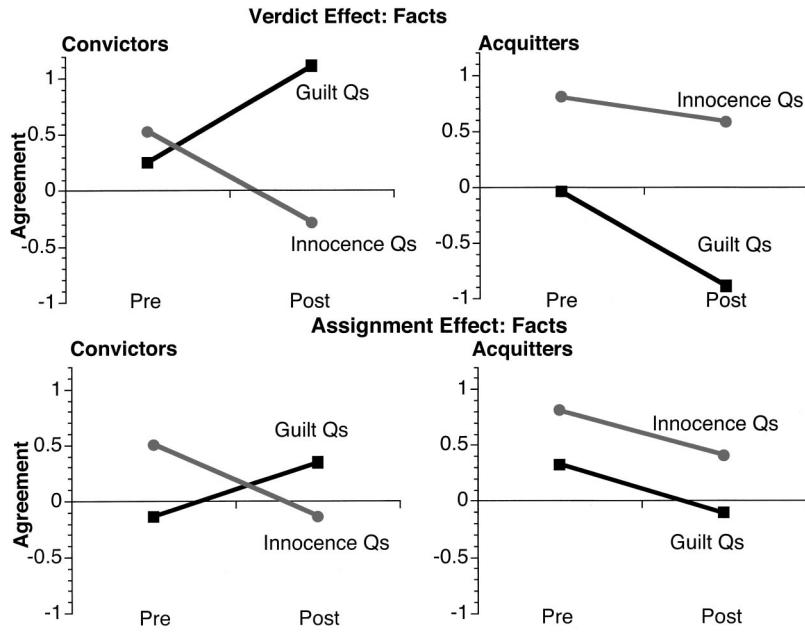


Figure 7. Study 4: Impact of both verdict and assignment to a side on evaluation of guilt and innocence facts by acquittees and convictors. Qs = questions; Pre = pretest; Post = posttest.

In order to do these analyses, the dependent variables of evaluation of the evidence and beliefs had to be recoded as follows. First, in the previous studies and in the above analyses for this study, we treated the guilt and the innocence items as separate repeated measures, with high scores on the guilt items being consistent with a guilty verdict and high scores on the innocence items being consistent with an innocent verdict. Here, we recoded

the innocence items so that higher scores were now consistent with guilt and then summed them with the guilt items to form a single guilt composite where higher scores were consistent with a guilty verdict. We did this separately for both the fact and the belief items. Second, we had previously treated the pretest and the posttest scores as separate repeated measures. Here we calculated a difference score between the pretest and the posttest scores on

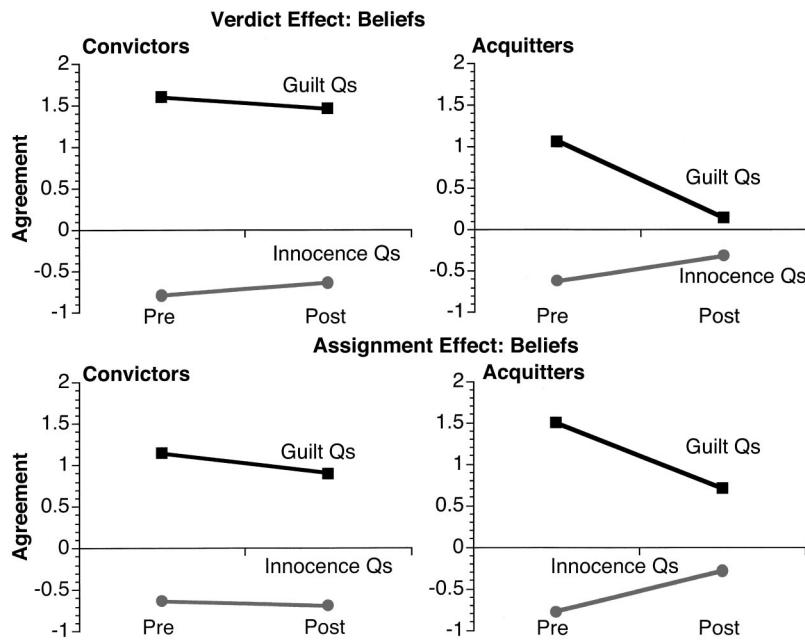


Figure 8. Study 4: Impact of both verdict and assignment to a side on evaluation of guilt and innocence beliefs by acquittees and convictors. Qs = questions; Pre = pretest; Post = posttest.

the guilt composite. A significant main effect of verdict on the resulting difference score is equivalent to the test of the interaction between the pre–post guilt scores and verdict. Thus, as a result of the recoding of the guilt and innocence items into a single guilt composite, a significant main effect of the verdict on the guilt difference score would be equivalent to the significant three-way interactions among verdict, pre–post, and guilt versus innocence items that provided evidence of coherence shifts in our previous analyses. We refer to this new score as the *coherence score*.

We first analyzed the fact questions, using linear regression to perform a mediational analysis, as outlined by R. M. Baron and Kenny (1986) and Shrout and Bolger (2002). (See Figure 9 for the path models.) With only the assignment to a side entered, there was a significant relationship between assignment to sides of the case and the coherence score, $r(161) = .215, p = .006$, such that being assigned to help the judge write an opinion supporting Jason's guilt led to greater endorsement of the facts in a direction consistent with guilt. To examine whether this effect was mediated by participants' verdicts, we first performed another regression in which we showed that assignment to a side, entered by itself, significantly predicted participants' verdicts, $r(161) = .215, p = .006$. (The identical value for these correlations is due to chance and is not a typographical error.) We then estimated the influence of verdict on the coherence score, controlling for the impact of assignment to side. This was significant, $\beta(160) = .406, p < .001$. In a final regression analysis, with verdict entered first, the impact of assignment to side on the coherence score was no longer significant, $\beta(160) = .128, p = .078$, suggesting that the verdict partially mediated the effect of assignment to side on the coherence score. To further corroborate this mediation, we directly tested the significance of the indirect path. Following Baron and Kenny's recommendations (see also Shrout & Bolger, 2002), we used Sobel's (1982) approximate significance test to obtain the indirect effect of the assignment to a side on the coherence score as mediated by the participants' verdict. This indirect effect was significant, $t(160) = 2.34, p < .05 (SE = 0.035)$. Thus, assignment to a side did influence participants' coherence shifts on the fact items, and this effect was at least partially mediated by participants' verdicts.

The analysis of the belief coherence scores was also consistent with this pattern, although weaker (see Figure 10). First, assignment to a side directly influenced the belief coherence score, $r(161) = .309, p < .001$. Second, assignment to a side directly

influenced participants' verdicts, $r(161) = .215, p = .006$. Third, verdict, controlling for assignment to a side, significantly influenced the belief coherence score, $\beta(160) = .210, p = .006$. Fourth, when we estimated the impact of assignment to a side on the belief coherence score, controlling for verdict, the effect was still significant, $\beta(160) = .264, p < .001$. However, the estimate of the indirect effect, using Sobel's (1982) test, was marginally significant, $t(160) = 1.92, p < .06 (SE = 0.024)$. Thus, assignment to sides of the case significantly influenced coherence shifts on the belief items as well, and there is some evidence that this effect was partially mediated by participants' verdicts.

This study provides further evidence for the bidirectionality of reasoning, and specifically for reasoning backwards from verdict to evidence. Not only did assignment to take a specific side significantly increase coherence shifts toward that side, for both the fact and belief items, but this effect was partially mediated by participants' verdicts.

General Discussion

The thrust of our findings is that the process of making decisions that are based on multiple judgments and inferences entails a reconstruction of the representation of the evidence. In all of the conditions, we observed coherence-driven shifts in the evaluations of both the facts and the underlying beliefs. When evaluated independently and spontaneously at the pretest phase, the evaluations of both the incriminating and exculpatory items received overall moderate agreement. In contrast, when virtually the same facts and beliefs were assessed in the context of the legal case, they were spread apart and represented in such a way that one decision dominated the other. These representations are said to be coherent in the sense that the evidence provided strong support for whichever verdict was chosen and weak or negative support for the rejected verdict. We found similar coherence shifts across significant changes in verdict distribution due to manipulations of the strength of evidence (Study 1) and differences in death penalty attitudes (Study 2). In Study 3, we found that when participants changed their minds in the course of making a decision, the reversal of the verdict was accompanied by a reversal of the evaluation of the evidence. Study 4 demonstrated that assignment of participants to verdicts caused the predicted shifts in the evaluation of the evidence. In sum, coherence shifts were found consistently across a number of different manipulations for differ-

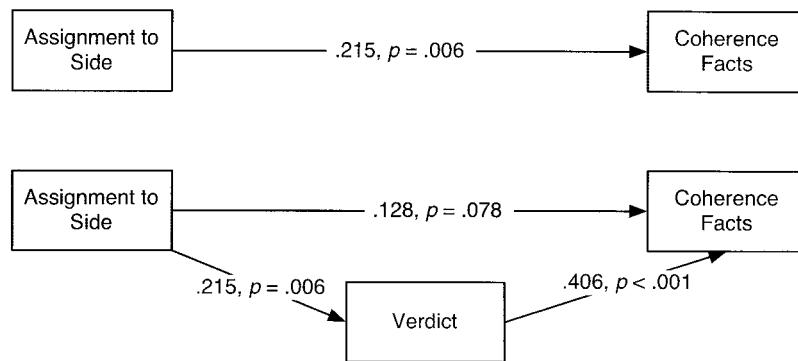


Figure 9. Study 4: Path models for mediational analysis for coherence facts.

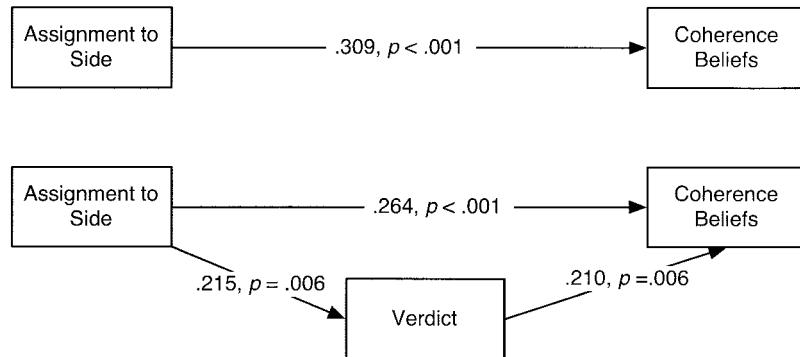


Figure 10. Study 4: Path models for mediational analysis for coherence beliefs.

ent kinds of crimes and with participants playing the three different roles of arbitrator, juror, and legal intern.

Another consistent finding was that the participants reported high levels of confidence in their decisions. High levels of confidence might seem at first remarkable given the fact that the decision task was ambiguous and complex. However, this finding is readily explainable in light of the fact that at the point of decision, the participants' mental representations of the task contained one decision alternative dominating the other. As expected, confidence levels were related to both the strength of the evidence at posttest (see Gill, Swann, & Silvera, 1998) and the magnitude of the coherence shifts.

We also found support for the bidirectional effect between evidence and verdict. Not only does the evaluation of the evidence influence the eventual verdict, but the developing verdict also seems to affect the evaluation of the evidence. The support for the latter direction provided by the first three studies is largely correlational, because participants chose their own verdict. Study 4 provides experimental evidence for the impact of the verdict on the evaluation of the evidence. In Study 4 we manipulated participants' assignment to a side in a legal case and showed that this assignment influenced both participants' verdicts and their evaluation of the evidence. Moreover, the impact of assignment to a side on the evaluation of the evidence was partially mediated by the participants' verdicts, indicating that not only does the evidence influence the verdict but that the verdict also influences the evaluation of the evidence.

In sum, these findings support our view that cognitive consistency theories—as actualized by parallel constraint satisfaction mechanisms in a connectionist network—can serve as a useful model for thinking about processing tasks that require combination of a multitude of probabilistic evidence into a discrete judgment. Gestaltian-induced forces serve to transform seemingly intractable tasks into coherent and manageable representations, thus enabling confident decisions and facilitating action. The reasoning process is bidirectional: Although the decisions were apparently based on the inferences made from the evidence, at the same time, the evaluations of the evidence were skewed so as to maintain coherence with the emerging decisions (Holyoak & Simon, 1999; Simon, 1998). The findings resemble research by Aliche, Davis, and Pezzo (1994) that showed that assessments of factual situations are influenced by the subjects' knowledge of the outcomes of those situations. The findings also resemble research by Asch (1940,

1946a) and Nisbett and Wilson (1977) on holistic judgments in person perception.

Coherence-driven processing appears to be an adaptive mechanism: It enables judgment, choice, and decision making in the face of complexity and ambiguity and thus makes it possible for people to confidently adopt a chosen course of action in what would otherwise be an intractable and daunting task. It must not be overlooked, however, that this functional and adaptive mechanism can entail a certain distortion of the evidence, in that coherent representations could not be attained without affecting change in the erstwhile inconsistent pattern of facts. There is some counter-normativity in the fact that the purported bases for a decision are affected by the decision they are supposed to justify (Simon, 1998, in press).

It should be noted that the consequences of coherence-driven decision making are of acute social importance. One notable example is the phenomenon of false convictions in criminal trials. Given the mounting revelations of false convictions (Forst, 2003) this inquiry serves more than mere theoretical interest. The mechanisms that enabled our participants to make confident decisions in the face of ambiguous evidence also make it more likely that in real life—especially when the decision maker experiences a strong need to produce unequivocal decisions—people will make decisions on the basis of insufficient evidence. Given the high levels of confidence that follow from parallel constraint satisfaction processing, it is not hard to see how weak and ambiguous evidence can be constructed to seem incriminating beyond a reasonable doubt (Simon, in press).

Parallel Constraint Satisfaction Processes

Our results are clearly consistent with computer simulations of parallel constraint satisfaction processing (Holyoak & Simon, 1999; Holyoak & Thagard, 1989; Read et al., 1997) as well as with experimental findings of constraint satisfaction (Holyoak & Simon, 1999; Read & Marcus-Newhall, 1993; Simon et al., 2001; Spellman et al., 1993; Spellman & Holyoak, 1992). The present research has provided the first findings of parallel constraint satisfaction processing in the making of decisions on the basis of probabilistic evidence.

As a number of people have noted, states of coherence or equilibrium in a constraint satisfaction network are “attractors,” so called because they effectively attract the state of the system to that

point. As first demonstrated by Hopfield (1982, 1984; see also Amit, 1989; Hertz et al., 1991; Rumelhart, Smolensky, McClelland, & Hinton, 1986), constraint satisfaction systems can be treated as if they have a certain energy, where the energy of the system is a function of the sum of the product of the activation of all possible pairs of nodes times the weight between them: $w_{ij} * a_i * a_j$, where a_i and a_j are the activations of the nodes and w_{ij} is the weight between them. Specifically, $E = -\sum \sum w_{ij} * a_i * a_j$. In this equation, the energy is minimized when the sign of the product of the activations is consistent with the sign of the weight between them, that is, the two activations are positive, and the weight between them is positive, or the two activations differ in sign, and the weight between them is negative.

This energy can be thought of as corresponding to the degree of coherence of the system. Higher coherence leads to lower energy. Hopfield (1982, 1984) showed that a constraint satisfaction system essentially acts to minimize its energy (in our terms, to maximize its consistency). He further noted that the energy of the system could be plotted as a surface in a multidimensional space, where the energy minimums look like valleys in the space. These valleys are attractors, because states of the system that fall on the slopes of the valley tend to move toward the bottom of the valley. That is, they act to attract states of the system that fall within these basins of attraction. Further, once the state of the system has moved into an attractor, the energy of the system (or its degree of inconsistency) must increase in order for the system to be able to move into another attractor.

One important further implication of this analysis is that coherence or parallel-constraint-based processing arises in any kind of physical system, such as the brain, that can be characterized by Hopfield's (1982, 1984) analysis. That is, coherence-based processing is not a special process or adaptation but instead arises from the fundamental nature of brain-based processing, which is based on the simultaneous interaction of numerous units (Rumelhart & McClelland, 1986).

Holistic Versus Mathematical Models of Cognition

The systematic change in the evaluations of the evidence over the course of making the decision poses a strong challenge to the descriptive validity of Anderson's (1981, 1996) IIT and to Bayes's theorem (see also Pennington & Hastie, 1992). As described above, these models make the assumptions of meaning invariance and value-integration independence. That is, they assume that the meaning of one attribute should not affect the meaning or value of another attribute and that the integration of information should not affect the underlying evaluation of that information. However, the observed coherence shifts provide strong evidence that both of these principles were violated by our participants. The shift of the evidence toward its ultimate lopsided state is flatly inconsistent with computations of weighted averages and with multiplication of probabilistic values of the pretest evaluations. A better explanation is that the numerous pieces of evidence interacted dynamically with the emerging verdict and possibly also with each other, leading to this lopsided, coherent representation of the case.

To some, this comparison of IIT (weighted attribute models) with holistic models might seem reminiscent of a previous theoretical duel in the context of impression formation between meaning invariance (e.g., Anderson, 1971; Kaplan, 1971, 1974) and

meaning change (Hamilton & Zanna, 1974; Wyer, 1974; Zanna & Hamilton, 1977). The former position maintained that impression formation could be explained on the basis of IIT, whereas the latter group followed Asch (1946a), and posited a gestaltlike model. That controversy came to a halt, though absent a clear resolution, following a critical review by Thomas Ostrom (1977). Two important differences should be noted between that previous research and our current experiments. First, the tasks in those experiments were mostly a cursory judgment of likableness, based only on a handful—often no more than two—trait adjectives; in contrast, our materials presented participants with a large, complex, and engaging task (see Devine & Ostrom, 1985). There is good reason to believe that this legal case was processed more extensively and in greater depth, thus leading to stronger representational effects (see Zajonc, 1960). Second, in the original research, the independent variables were manipulated orthogonally, and much of the disagreement revolved around the interpretations of the results (for a review, see Ostrom, 1977). Our proof of change is more direct in that we measured evaluative shifts directly, that is, by comparing two measurements of virtually the same evidence.

Our findings share some features with Roe, Busemeyer, and Townsend's (2001) multialternative decision field theory, which relies on a recurrent connectionist network to make a choice among alternatives. However, in their model, the recurrent network only applies to alternatives for which the valences have already been calculated. That is, the network represents the relationships among the overall valence of the different alternatives but does not represent the relationships among the attributes that underlie those valences. Thus, the evaluation of the different alternatives by the recurrent network cannot affect the evaluation of the attributes of the different choices. In contrast, we have shown that the individual's choices (verdicts) do influence the evaluations of the pieces of evidence (attributes).

Our research also sheds light on a recurrent finding that has baffled legal psychologists. In experimentation on Bayesian modeling in legal decision making, researchers have been surprised by the observed interdependence between evaluations of seemingly unrelated pieces of evidence. For example, B. C. Smith, Penrod, Otto, and Park (1996) found increases in ratings of prosecution evidence after the presentation of the defendant's weak alibi evidence. Similarly, Schum and Martin (1982) found it "striking" that inconsistencies among pieces of evidence were interpreted as consistencies and that redundancies were "double counted" and interpreted as bolstering one another (p. 144). These phenomena were observed only when participants were instructed to make global judgments of the evidence; they were not obtained when the evidence was assessed piecemeal. All of these findings can, of course, be understood as coherence effects that result from constraint satisfaction processing.

Relationship With Pennington and Hastie's Story Model

Our findings are consistent with many of the central features of Pennington and Hastie's (1986, 1988, 1993) story model. Our findings support the view that decisions are determined by constructions of representations rather than by the "raw evidence" (Pennington & Hastie, 1988, p. 523), that the emerging story serves to direct interpretation of the evidence (Pennington & Hastie, 1986, 1993), and that confidence in the decision is a

function of the coherence of the reconstruction (Pennington & Hastie, 1988). It is also noteworthy that the findings of Study 3 support Pennington and Hastie's conclusion that representational changes are not merely post hoc justifications of the verdict; rather, they occur before the decision maker has committed herself to a final conclusion (Pennington & Hastie, 1988, 1993).

Our findings thus provide support for the central features of the story model, which, to the best of our knowledge, has not heretofore been replicated. Our findings also add some strength to the story model. Although the data supporting the model are compelling, it should be noted that the methodology used to prove it did not include a baseline measurement of the evidence. Our design offers a more direct measurement of the reconstruction of the evidence by comparing the participants' evaluation of the evidence with their own original representation of it.

At the same time, there are important ways in which the story model is limited compared with the coherence-based model we outline. As encapsulated in its name, the story model proposes that evidence is constructed in the form of stories that bear a distinct narrative structural order. Specifically, "stories involve human action sequences connected by relationships of physical causality and intentional causality between events" (Pennington & Hastie, 1993, p. 196; see also Pennington & Hastie, 1986, 1988). The case on which the theory was based, *Commonwealth of Massachusetts v. Johnson*, deals with a murder charge that hinged on a determination of the defendant's mental state. The prosecution charged that the homicide was committed with premeditation, whereas the defense denied any such malice and argued instead for self-defense. Indeed, the case contained a great deal of ambiguity regarding Johnson's mental state, and the participants' reasoning was replete with inferences about intentions, psychological states, goals, and motivations—the intentional and causal material that glues episodes into a globally coherent story of "what happened." For instances of this kind, the story model's insights are undoubtedly relevant and valuable. Its relevance for some other kinds of legal and nonlegal decision tasks, however, is somewhat limited. Many legal cases do not involve intentional and causal schemes, for example, where the issue in question hinges on the reasonability of one's risky behavior or on the legality of one's physical actions. The model's applicability is questionable also in tasks of evidence integration outside the realm of legal decisions, such as in developing scientific explanations or making judgments about physical events (cf. Thagard, 1989).

In contrast, the research on coherence-based processing by constraint satisfaction has a broader scope of application. As mentioned above, it has been shown to explain a wide array of tasks, including social reasoning (Read & Marcus-Newhall, 1993), analogical reasoning (Spellman & Holyoak, 1992), attitude change (Spellman et al., 1993), multiattribute decision making (Simon et al., in press), and abstract legal reasoning (Holyoak & Simon, 1999; Simon et al., 2001). The evidentiary question in the case of Jason Wells revolves exclusively around identifying the culprit. The evidence was explicitly designed to be physical, circumstantial, and disjointed and thus not susceptible to being explained by overarching causal and intentional schemes or by any other apparent narrative structure (except of course, a tautological and reductive schema of the sort "the defendant is guilty" or "the defendant is innocent"). Thus, our findings suggest that the restructuring of cognitive representations is not limited to instances in which the

evidence lends itself to narrative structures revolving around themes of human intentionality and causation. Coherence shifts are best conceived as a basic and general tendency of the cognitive system to drift away from incoherent representations toward coherent states of equilibrium. The story model can be understood as a specific form of this general tendency. Although there is good reason to believe that coherence shifts are particularly pronounced when the task involves schemes of human intentionality, these schemes do not appear to be essential.

The Predecisional Nature of the Coherence Shifts

Study 3 provides support for the contention that the coherence shifts occur prior to the final commitment to a decision (Simon et al., 2001, in press). We believe that parallel constraint satisfaction processes play an active role in the decision-making process in that they spread apart the subsets of evidence and thus create a state in which one verdict dominates the other (Montgomery, 1983; Svenson, 1992). One might object to this conclusion by interpreting our findings as mere postdecisional dissonance reduction: a process of allaying the sense of regret caused by the postdecisional salience of the negative features of the chosen decision alternative and the positive features of the rejected alternative (Festinger, 1957, 1964). According to this approach, our findings of coherence shifts could be explained not as part of the decision-making process at all but rather as the reduction of dissonance aroused by preliminary decisions made prior to our measurement of evidence evaluation. This objection raises a number of weighty theoretical issues warranting a discussion that extends beyond the scope of the current article. Short of a complete discussion, we suggest in brief that our conclusion is defensible on two grounds.

First, there is mounting empirical evidence of extensive predecisional processing that is consistent with our view (e.g., Boiney, Kennedy, & Nye, 1997; Brownstein, Read, & Simon, in press; Montgomery, Selart, Garling, & Lindberg, 1994; Pennington & Hastie, 1988, 1993; Phillips, 2002; Svenson, 1992). For a comprehensive review, see Brownstein (2003). Specifically, our findings are consistent with the findings by Russo and his colleagues of predecisional "distortions" in processing of information (e.g., Carlson & Russo, 2001; Russo, Medvec, & Meloy, 1996). In their experiments, participants were presented with one attribute at a time. Immediately after each attribute was presented, participants were asked to evaluate that attribute and to make an overall evaluation of the decision based on the available information. Participants are found to interpret each piece of incoming evidence in a way that is biased toward their preferred verdict. It is theoretically possible that these repeated intervening questions bear some effect on the responses elicited. In contrast, our research is designed to encourage participants to process the evidence globally as they would normally do rather than incrementally. The fact that similar findings are obtained by means of these two different methodologies is a testament to the robustness of the psychological phenomenon.

Second, the objection seems to overlook cognitive dissonance's emphasis on the necessity of commitment to the arousal of dissonance. Festinger explained, "We must accept the fact that dissonance-reduction processes do not automatically start when a decision is made. . . . simply making a decision is not enough. The decision must have the effect of committing the person" (Fest-

inger, 1964, p. 42). The centrality of commitment to the onset of dissonance is similarly emphasized in the theoretical work of Brehm and Cohen (1962), Kiesler (1971), and Wicklund and Brehm (1976) and in the experimentation of Gollwitzer and his colleagues (Beckmann & Gollwitzer, 1987; Gollwitzer & Bayer, 1999).

We agree that the moment of commitment is of psychological significance and that it is an appropriate criterion for distinguishing between pre- and postdecisional phases of processing. Our contention is that the shifts observed at the interim phase of Study 3 occurred prior to the participants' commitment to any verdict. That is immediately apparent from the fact that a substantial minority did in fact switch their decision after that point, and they did so in spite of a conceivable experimental demand to maintain consistency (Aronson, 1968).

The design of Study 3 also captures an important dimension of external validity, in that many of the most important decisions in people's lives require extensive and prolonged processing, especially when the choice is a difficult one. It is not uncommon for decision makers to spend days, weeks, and months pondering decisions, nor is it uncommon that one's leanings can shift from one alternative to the other, with the vacillation continuing into the eleventh hour (Janis & Mann, 1977). There is strong empirical evidence that people process information and form opinions without any instruction to do so (Carlston & Skowronski, 1994; Simon et al., 2001; Winter & Uleman, 1984), even when the text is presented as a distractor to some other ostensible processing task (Winter, Uleman, & Cunniff, 1985). It is probably the case that during a trial, jurors assess their leanings many times, and those who are genuinely open minded switch their leanings back and forth between the verdicts. In sum, it is probably the case that good decision makers continuously test and retest the mental models of their decisions before arriving at a final conclusion. To say that the prolonged, thorough, multistaged real-life processes are no more than postdecisional processing (except for the very first tentative iteration) offers a reductionist view of human cognition that effectively empties the crucial predecisional processing of any significance.

Changes in Background Beliefs and Knowledge

The second objective of this research was to test whether coherence effects also affect the way background beliefs are accessed and constructed. Our findings support this hypothesis. Throughout the studies, the general beliefs reported by our participants in the context of the legal case were systematically different from those they reported when asked in the context of isolated vignettes. Again, these changes were in the direction of increased coherence with the evaluation of the related pieces of evidence and with the verdict chosen by that participant. It is not surprising that the observed changes in the background beliefs are somewhat less consistent and weaker than the changes in the evaluations of the facts. In contrast to the pieces of evidence that were specific to this case, general beliefs are typically embedded in larger knowledge systems and attitudinal structures and should thus be expected to be more resistant to change.

The observed changes in beliefs are consistent with the findings of Klein and Kunda (1992). They manipulated the participants' goal of the impression formation (expecting to interact closely

with a schizophrenic person or not) and found that those who expected to interact with the target person reported more favorable attitudes toward schizophrenic people than did participants who were not led to expect an interaction. Our findings differ from that research in that we did not manipulate the evidence that was related to the beliefs. The variables that we manipulated were unrelated to the beliefs in question—other pieces of evidence (Study 1) and the new evidence in Study 3—yet the beliefs changed in the predicted manner. In Study 2, participants' death penalty attitudes were also unrelated to the beliefs in question. Yet participants with different attitudes tended to arrive at different verdicts, and their beliefs shifted to be consistent with those verdicts. Our interpretation of the observed changes in beliefs is that coherence-driven structural forces imposed coherence not only on the pieces of evidence themselves but also on the background knowledge on which they were supposedly based. It appears, then, that the reaches of connectionist influences are extensive.

Redux of Cognitive Consistency Theories

Our results manifest the gestaltian principles that underlie cognitive consistency theories (Heider, 1946, 1958). As would be predicted by a model of structural dynamics, the representation of the initially inconsistent set of facts drifted toward a point at which their interaction with one another and with the chosen verdict equilibrated at a configurational state of good form. The view that coherence-driven processing enables confident choice is consistent with the views expressed by some consistency theorists. For example, Percy Tannenbaum (1968b) argued that

the reasoning behind [the consistency position] relates to the organism's presumed need to apprehend and comprehend things and events about him. In monitoring, processing, and interpreting information from the environment, some degree of consistency and equilibrium is seen as essential for reasons of parsimony and economy of effort, as well as to allow for the predictability of, and hence adaptability to, subsequent encounters. . . . most assume a universal value for the organism in his having a stable predictable view of his environment. (pp. 343–344)

Similarly, Abelson (1968) described the consistency function as intended to "organize the information stored by the individual in a way that is likely to be useful to him, directly or indirectly, for affective or behavioral purposes" (p. 133). Heider (1979) suggested that the tendency toward balanced states is based on the desire "to have our cognitive food prepared so that it is easy to swallow, to assimilate" (p. 16). Albert Pepitone (1966) explained that consistent structures "are simpler to maintain than distinctions, discrepancies and contradictions" (p. 270; see also Back, 1968).

It is particularly noteworthy that the observed findings also fit nicely with some of the more insightful and sophisticated models offered by consistency theorists. One example is William McGuire's (1960) "loose-link" network built from extended chains of reasoning. As McGuire conceived the reasoning process, consistency-driven forces spread through the related links to the system's remote elements, gradually changing the network's activation toward a higher state of coherence (McGuire, 1960). Similarly, Abelson and Rosenberg's (1958) theory of symbolic psycho-logic (see also Rosenberg & Abelson, 1960), carves out a

conceptual arena that contains the numerous concepts and propositions that make up the judgment task. The myriad of interactions within this arena determines the structural properties of the task. States of imbalance are driven toward equilibrium by means of changing the represented elements. In this vein, as Abelson and Rosenberg explained, task variables are "manipulated in thought" (Abelson & Rosenberg, 1958, p. 4). Our findings are also consistent with Judson Mills's choice certainty theory (Mills, 1965, 1968), according to which people tend to make decisions with sufficiently high levels of certainty. The absence of certainty triggers cognitive processes that search and interpret information in ways that make the emerging choice seem superior to its alternative (Mills, 1965, 1968). Connectionist representations and constraint satisfaction processing mechanisms were presciently anticipated in Milton Rosenberg's (1968) metaphoric "attitudinal cognitorium." Rosenberg imagined a large physical space in which all the variables are represented by metal disks. The disks are interconnected; consistent disks are linked by green strings and inconsistent ones by red strings. To obtain the result of a complex attitudinal position, one must pull the disks to the third dimension and thus enable the structural interconnections to settle the system at a stable state (Rosenberg, 1968, pp. 78–80).

The current research lends support to the view that structural dynamics is an appropriate model for human cognition, as captured by classical formulations of consistency theories and subsequently embodied in parallel constraint satisfaction processing (Read et al., 1997; Read & Miller, 1994). These experiments underscore the ability of consistency-based models of cognition to handle realistic, complex mental processes. We suggest that with the support of recent advances in connectionist representations and parallel constraint satisfaction processing, cognitive consistency theories might yet achieve their unfulfilled promise and come closer to the aspirations shared by Lewin (1935), Abelson (1968), McGuire (1968), and others to develop a general "psychology of inference" (McGuire, 1968, p. 140).

References

- Abelson, R. P. (1968). Psychological implication. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & R. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 112–139). Chicago: Rand McNally.
- Abelson, R. P., & Rosenberg, M. (1958). Symbolic psycho-logic: A model of attitudinal cognition. *Behavioral Science*, 3, 1–8.
- Alicke, M. D., Davis, T. L., & Pezzo, M. V. (1994). A posteriori adjustment of a priori decision criteria. *Social Cognition*, 12, 281–308.
- Amit, D. J. (1989). *Modeling brain function: The world of attractor neural networks*. Cambridge, England: Cambridge University Press.
- Anderson, N. H. (1962, November 16). Application of an additive model to impression formation. *Science*, 138, 817–818.
- Anderson, N. H. (1968). A simple model for information integration. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & R. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 731–743). Chicago: Rand McNally.
- Anderson, N. H. (1971). Two more tests against change of meaning in adjective combinations. *Journal of Verbal Learning and Verbal Behavior*, 10, 75–85.
- Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.
- Anderson, N. H. (1996). *A functional theory of cognition*. Mahwah, NJ: Erlbaum.
- Aronson, E. (1968). Dissonance theory: Progress and problems. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & R. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 5–27). Chicago: Rand McNally.
- Asch, S. E. (1940). Studies in the principles of judgments and attitudes: II. Determination of judgments by group and by ego standards. *Journal of Social Psychology*, 12, 433–465.
- Asch, S. E. (1946a). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 258–290.
- Asch, S. E. (1946b). Max Wertheimer's contribution to psychology. *Social Research*, 13, 81–102.
- Back, K. W. (1968). Equilibrium as motivation: Between pleasure and enjoyment. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & P. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 311–318). Chicago: Rand McNally.
- Baron, J. (2000). *Thinking and deciding* (3rd ed.). New York: Cambridge University Press.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Beckmann, J., & Gollwitzer, P. M. (1987). Deliberative versus implemental states of mind: The issue of impartiality in predecisional and post-decisional information processing. *Social Cognition*, 5, 259–279.
- Boiney, L. G., Kennedy, J., & Nye, P. (1997). Instrumental bias in motivated reasoning: More when more is needed. *Organizational Behavior and Human Decision Processes*, 72, 1–24.
- Brehm, J. W., & Cohen, A. R. (1962). *Explorations in cognitive dissonance*. New York: Wiley.
- Brownstein, A. (2003). Biased predecision processing. *Psychological Bulletin*, 129, 545–568.
- Brownstein, A., Read, S. J., & Simon, D. (in press). Effects of individual expertise and task importance on pre-decision reevaluation of alternatives. *Personality and Social Psychology Bulletin*.
- Carlson, K. A., & Russo, J. E. (2001). Biased interpretation of evidence by mock jurors. *Journal of Experimental Psychology: Applied*, 7, 91–103.
- Carlston, D. E., & Skowronski, J. J. (1994). Savings in the relearning of trait information as evidence for spontaneous inference generation. *Journal of Personality and Social Psychology*, 66, 840–856.
- Cartwright, D., & Harary, F. (1956). Structural balance: A generalization of Heider's theory. *Psychological Review*, 63, 277–293.
- Cowan, C. L., Thompson, W. C., & Ellsworth, P. C. (1984). The effects of death qualification on jurors' predisposition to convict and on the quality of deliberation. *Law and Human Behavior*, 8, 53–79.
- Devine, P. G., & Ostrom, T. M. (1985). Cognitive mediation of inconsistency discounting. *Journal of Personality and Social Psychology*, 49, 5–21.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.
- Festinger, L. (1964). *Conflict, decision, and dissonance*. Stanford, CA: Stanford University Press.
- Friedman, R. D. (1997). Answering the Bayesioskeptical challenge. *International Journal of Evidence and Proof*, 1, 276–291.
- Forst, B. (2003). *Errors of justice: Nature, sources and remedies*. New York: Cambridge University Press.
- Gill, M. J., Swann, W. B., Jr., & Silvera, D. H. (1998). On the genesis of confidence. *Journal of Personality and Social Psychology*, 75, 1101–1114.
- Gollwitzer, P. M., & Bayer, U. (1999). Deliberative versus implementational mindsets in the control of action. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 403–422). New York: Guilford Press.

- Golob, H. F. (1974). The subject–verb–object approach to social cognition. *Psychological Review*, 81, 286–321.
- Goodman-Delahunty, J., Greene, E., & Hsiao, W. (1998). Construing motive in videotaped killings: The role of jurors' attitudes toward the death penalty. *Law and Human Behavior*, 22, 257–271.
- Hamilton, D. L., & Zanna, M. P. (1974). Context effects in impression formation: Changes in connotative meaning. *Journal of Personality and Social Psychology*, 29, 649–654.
- Harmon-Jones, E., & Mills, J. (Eds.). (1999). *Cognitive dissonance: Progress on a pivotal theory in social psychology*. Washington, DC: American Psychological Association.
- Hastie, R. (1993). Algebraic models of juror decision processes. In R. Hastie (Ed.), *Inside the juror: The psychology of juror decision making* (pp. 84–115). New York: Cambridge University Press.
- Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology*, 21, 107–111.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Heider, F. (1960). The Gestalt theory of motivation. In M. R. Jones (Ed.), *Nebraska Symposium on Motivation* (Vol. 8, pp. 145–172). Lincoln: University of Nebraska Press.
- Heider, F. (1979). On balance and attribution. In P. W. Holland & S. Leinhardt (Eds.), *Perspectives on social network research* (pp. 11–23). New York: Academic Press.
- Hertz, J., Krogh, A., & Palmer, R. G. (1991). *Introduction to the theory of neural computation*. Redwood City, CA: Addison-Wesley.
- Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General*, 128, 3–31.
- Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13, 295–355.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, USA*, 79, 2554–2558.
- Hopfield, J. J. (1984). Neurons with graded responses have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences, USA*, 81, 3088–3092.
- Insko, C. A. (1984). Balance theory, the Jordan paradigm, and the Wiest tetrahedron. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17, pp. 89–140). New York: Academic Press.
- Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. New York: Free Press.
- Kaplan, M. F. (1971). Context effects in impression formation: The weighted average versus the meaning–change formulation. *Journal of Personality and Social Psychology*, 19, 92–99.
- Kaplan, M. F. (1974). Context-induced shifts in personality trait evaluation: A comment on the evaluative halo effect and meaning change interpretations. *Psychological Bulletin*, 81, 891–895.
- Kaplan, M. F. (1975). Informational integration in social judgment: Interaction of judge and informational components. In M. F. Kaplan & S. Schwartz (Eds.), *Human judgment and decision processes* (pp. 139–171). New York: Academic Press.
- Kaplan, M. F., & Miller, L. E. (1978). Reducing the effects of juror bias. *Journal of Personality and Social Psychology*, 36, 1443–1455.
- Kaye, D. H. (2000). Bayes, burdens, and base rates. *International Journal of Evidence and Proof*, 4, 260–267.
- Kelman, H. C., & Baron, R. M. (1968). Determinants of modes of resolving inconsistency dilemmas: A functional analysis. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & R. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 670–683). Chicago: Rand McNally.
- Kerr, N. L. (1978). Severity of prescribed penalty and mock jurors' verdicts. *Journal of Personality and Social Psychology*, 36, 1431–1442.
- Kiesler, C. A. (1971). *The psychology of commitment: Experiments linking behavior to belief*. New York: Academic Press.
- Klein, W. M., & Kunda, Z. (1992). Motivated person perception: Constructing justifications for desired beliefs. *Journal of Experimental Social Psychology*, 28, 145–168.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480–498.
- Lempert, R. O. (1977). Modeling relevance. *Michigan Law Review*, 75, 1021–1057.
- Lewin, K. (1935). *A dynamic theory of personality*. New York: McGraw-Hill.
- Markus, H., & Zajonc, R. B. (1985). The cognitive perspective in social psychology. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 1, pp. 137–230). New York: Random House.
- McClelland, J. L., & Rumelhart, D. E. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol. 2. *Psychological and biological models*. Cambridge, MA: MIT Press.
- McGuire, W. J. (1960). Cognitive consistency and attitude change. *Journal of Abnormal and Social Psychology*, 60, 345–353.
- McGuire, W. J. (1968). Theory of the structure of human thought. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & R. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 140–162). Chicago: Rand McNally.
- Mills, J. (1965). The effect of certainty on exposure to information prior to commitment. *Journal of Experimental Social Psychology*, 1, 348–355.
- Mills, J. (1968). Interest in supporting and discrepant information. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & R. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 771–776). Chicago: Rand McNally.
- Montgomery, H. (1983). Decision rules and the search for a dominance structure: Towards a process model of decision making. In P. Humphreys, O. Svenson, & A. Vari (Eds.), *Analysis and aiding decision processes* (pp. 343–369). Amsterdam: North-Holland.
- Montgomery, H., Selart, M., Garling, T., & Lindberg, E. (1994). The judgment–choice discrepancy: Noncompatibility or restructuring? *Journal of Behavioral Decision Making*, 7, 145–155.
- Newcomb, T. M. (1953). An approach to the study of communicative acts. *Psychological Review*, 60, 393–404.
- Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 35, 250–256.
- O'Neil, K. M., Patry, M. W., & Penrod, S. D. (in press). Exploring the effects of attitudes toward the death penalty on capital sentencing verdicts. *Psychology, Public Policy, and Law*.
- Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review*, 62, 42–55.
- Ostrom, T. M. (1977). Between-theory and within-theory conflict in explaining context effects in impression formation. *Journal of Experimental Social Psychology*, 13, 492–503.
- Ostrom, T. M., Werner, C., & Saks, M. J. (1978). An integration theory analysis of jurors' presumptions of guilt or innocence. *Journal of Personality and Social Psychology*, 36, 436–450.
- Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology*, 51, 242–258.
- Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 521–533.
- Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the story model for juror decision making. *Journal of Personality and Social Psychology*, 62, 189–206.
- Pennington, N., & Hastie, R. (1993). The story model for juror decision making. In R. Hastie (Ed.), *Inside the juror: The psychology of juror decision making* (pp. 192–221). New York: Cambridge University Press.

- Pennington, N., & Hastie, R. (1994). Reasoning in explanation-based decision making. In P. N. Johnson-Laird & E. Shafir (Eds.), *Reasoning and decision making* (pp. 123–163). Cambridge, MA: Blackwell.
- Pepitone, A. (1966). Some conceptual and empirical problems of consistency models. In S. Feldman (Ed.), *Cognitive consistency: Motivational antecedents and behavior consequents* (pp. 257–297). New York: Academic Press.
- Phillips, F. (2002). The distortion of criteria after decision-making. *Organizational Behavior and Human Decision Processes*, 88, 769–784.
- Read, S. J. (1987). Constructing causal scenarios: A knowledge structure approach to causal reasoning. *Journal of Personality and Social Psychology*, 52, 288–302.
- Read, S. J., & Marcus-Newhall, A. (1993). Explanatory coherence in social explanations: A parallel distributed processing account. *Journal of Personality and Social Psychology*, 65, 429–447.
- Read, S. J., & Miller, L. C. (1994). Dissonance and balance in belief systems: The promise of parallel constraint satisfaction processes and connectionist modeling approaches. In R. C. Schank & E. J. Langer (Eds.), *Beliefs, reasoning, and decision making: Psycho-logic in honor of Bob Abelson* (pp. 209–235). Hillsdale, NJ: Erlbaum.
- Read, S. J., Snow, C. J., & Simon, D. (2003). Constraint satisfaction processes in social reasoning. In R. Alterman & D. Kirsch (Eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society* [CD-ROM]. Mahwah, NJ: Erlbaum.
- Read, S. J., Vanman, E. J., & Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and gestalt principles: (Re)introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review*, 1, 26–53.
- Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, 108, 370–392.
- Rosenberg, M. J. (1960). An analysis of affective–cognitive consistency. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 15–64). New Haven, CT: Yale University Press.
- Rosenberg, M. J. (1968). Hedonism, inauthenticity, and other goads towards expansion of a consistency theory. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & R. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 73–111). Chicago: Rand McNally.
- Rosenberg, M. J., & Abelson, R. P. (1960). An analysis of cognitive balancing. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 112–163). New Haven, CT: Yale University Press.
- Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol. 1. Foundations. Cambridge, MA: MIT Press.
- Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models* (pp. 7–57). Cambridge, MA: MIT Press.
- Russo, E. J., Medvec, V. H., & Meloy, M. G. (1996). The distortion of information during decisions. *Organizational Behavior and Human Decision Processes*, 66, 102–110.
- Schank, R. C., & Abelson, R. P. (1995). Knowledge and memory: The real story. In R. S. Wyer Jr. (Ed.), *Knowledge and memory: The real story* (pp. 1–85). Hillsdale, NJ: Erlbaum.
- Schum, D. A. (1994). *The evidential foundations of probabilistic reasoning*. New York: Wiley.
- Schum, D. A., & Martin, A. W. (1982). Formal and empirical research on cascaded inference in jurisprudence. *Law and Society Review*, 17, 105–151.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and non-experimental studies: New procedures and recommendations. *Psychological Methods*, 7, 422–445.
- Shultz, T. R., & Lepper, M. R. (1996). Cognitive dissonance reduction as constraint satisfaction. *Psychological Review*, 103, 219–240.
- Shultz, T. R., & Lepper, M. R. (1998). The consonance model of dissonance reduction. In S. J. Read & L. C. Miller (Eds.), *Connectionist models of social reasoning and social behavior* (pp. 211–244). Mahwah, NJ: Erlbaum.
- Simon, D. (1998). A psychological model of judicial decision making. *Rutgers Law Journal*, 30, 1–142.
- Simon, D. (in press). A third view of the black box: Cognitive coherence in legal decision making. *University of Chicago Law Review*.
- Simon, D., & Holyoak, K. J. (2002). Structural dynamics of cognition: From consistency theories to constraint satisfaction. *Personality and Social Psychology Review*, 6, 283–294.
- Simon, D., Krawczyk, D. C., & Holyoak, K. J. (in press). Construction of preferences by constraint satisfaction. *Psychological Science*.
- Simon, D., Pham, L. B., Le, Q. A., & Holyoak, K. J. (2001). The emergence of coherence over the course of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1250–1260.
- Simon, D., Snow, C. J., Brownstein, A. L., & Read, S. J. (2002, March). *Testing juror decision making over the Internet*. Paper presented at the Annual Conference of the American Psychology–Law Society, Austin, TX.
- Smith, B. C., Penrod, S. D., Otto, A. L., & Park, R. C. (1996). Jurors' use of probabilistic evidence. *Law and Human Behavior*, 20, 49–82.
- Smith, E. R. (1996). What do connectionism and social psychology have to offer each other? *Journal of Personality and Social Psychology*, 70, 893–912.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhart (Ed.), *Sociological methodology* (pp. 290–312). San Francisco: Jossey-Bass.
- Spellman, B. A., & Holyoak, K. J. (1992). If Saddam is Hitler then who is George Bush? Analogical mapping between systems of social roles. *Journal of Personality and Social Psychology*, 62, 913–933.
- Spellman, B. A., Ullman, J. B., & Holyoak, K. J. (1993). A coherence model of cognitive consistency: Dynamics of attitude change during the Persian Gulf War. *Journal of Social Issues*, 49, 147–165.
- Svenson, O. (1992). Differentiation and consolidation theory of human decision making: A frame of reference for the study of pre- and post-decision processes. *Acta Psychologica*, 80, 143–168.
- Tannenbaum, P. H. (1968a). The congruity principle: Retrospective reflections and recent research. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & R. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 52–72). Chicago: Rand McNally.
- Tannenbaum, P. H. (1968b). Is there anything special about consistency? In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & R. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 341–346). Chicago: Rand McNally.
- Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, 12, 435–467.
- Wertheimer, M. (1967a). The general theoretical situation. In W. D. Ellis (Ed.), *A source book of Gestalt theory* (pp. 12–16). New York: Humanities Press. (Original work published 1922)
- Wertheimer, M. (1967b). Gestalt theory. In W. D. Ellis (Ed.), *A source book of Gestalt theory* (pp. 1–11). New York: Humanities Press. (Original work published 1924)
- Wertheimer, M. (1967c). Laws of organization in perceptual forms. In

- W. D. Ellis (Ed.), *A source book of Gestalt theory* (pp. 71–88). New York: Humanities Press. (Original work published 1923)
- Wicklund, R. A., & Brehm, J. W. (1976). *Perspectives on cognitive dissonance*. Hillsdale, NJ: Erlbaum.
- Winter, L. & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology*, 47, 237–252.
- Winter, L., Uleman, J. S., & Cunniff, C. (1985). How automatic are social judgments? *Journal of Personality and Social Psychology*, 49, 904–917.
- Wyer, R. S. (1974). Changes in meaning and halo effects in personality impression formation. *Journal of Personality and Social Psychology*, 29, 829–835.
- Wyer, R. S., & Radvansky, G. A. (1999). The comprehension and validation of social information. *Psychological Review*, 106, 89–118.
- Zajonc, R. B. (1960). The process of cognitive tuning in communication. *Journal of Abnormal and Social Psychology*, 61, 159–167.
- Zajonc, R. B. (1968). Cognitive theories in social psychology. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (2nd ed., Vol. 1, pp. 320–411). Reading, MA: Addison-Wesley.
- Zajonc, R. B. (1983). Discussion of Abelson's talk on Cartwright's founder's day. *Personality and Social Psychology Bulletin*, 9, 55–59.
- Zanna, M. P., & Hamilton, D. L. (1977). Further evidence for meaning change in impression formation. *Journal of Experimental Social Psychology*, 13, 224–238.

Appendix

Posttest Questions About Evidence in the Testimony and Related Beliefs

- 1a. If the defendant had committed the crime, it is unlikely that he could have made it in time for the 8 PM swim meet. (IF)
- 1b. When driving in evening traffic, no matter how aggressively one drives, it is very difficult to shorten the travel time substantially. (IB)
- 2a. The defendant repaid his debt to the credit card company with the money he received from his sister Lynn. (IF)
- 2b. In the flower business, financial deals are typically done in cash. (IB)
- 3a. The technician's identification of the defendant makes it likely that the person hurrying out of the bookkeeper's office was in fact the defendant. (GF)
- 3b. In general, when people identify someone whom they've already seen once or twice before the identifications are accurate. (GB)
4. The fact that the car seen leaving the scene was the same as the defendant's makes it likely that it was the defendant who was captured on the video camera driving away. (GF)
5. Generally speaking, people who have committed a crime, are likely to commit additional crimes somewhere down the line. (GB)
- 6a. The defendant was looking for a way to get back at the company for the disciplinary measures taken against him. (GF)
- 6b. Generally speaking, when people feel that they have been treated unjustly, they often do something spiteful to get back. (GB)
- 6c. The defendant's reaction to the unfair disciplinary action was to prove himself to his boss by working even harder. (IF)
- 6d. Generally speaking, when people feel that they have been treated unjustly, they typically try to prove themselves by doing even better at their job. (IB)
- 7a. *The information provided by the cell phone company regarding the defendant's whereabouts makes it likely that he was at the crime scene. (GF)
- 7b. *The cell phone company's ability to determine locations of callers is extremely accurate. (GB)
- Note.* G = guilt item; I = innocence item; F = fact item; B = general world belief.
- *Questions 7a and 7b were not included in Study 1.

Received November 18, 2002

Revision received January 27, 2004

Accepted February 3, 2004 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://watson.apa.org/notify/> and you will be notified by e-mail when issues of interest to you become available!