Journal of Experimental Psychology: General
1999, Vol. 128, No. 1, 3-31

0096-3445/99/$3.00

Bidirectional Reasoning in Decision Making by Constraint Satisfaction

Keith J. Holyoak
University of California, Los Angeles

Dan Simon
Haifa University

Recent constraint satisfaction models of explanation, analogy, and decision making claim that these processes are influenced by bidirectional constraints that promote coherence. College students were asked to reach a verdict in a complex legal case involving multiple conflicting arguments, including alternative analogies to the target case. Participants rated agreement with the individual arguments both in isolation before seeing the case and again after reaching a verdict. Assessments of the individual arguments shifted so as to cohere with their emerging verdict. A cascade of spreading coherence influenced decisions made about a subsequent case involving different legal issues. Participants' memory for their initial positions also shifted so as to cohere with their final positions. The results demonstrate that constraint satisfaction can transform ambiguous inputs into coherent decisions.

One of the most deep-rooted assumptions about human reasoning is that the flow of inference is inherently unidirectional, moving from premises to be accepted as given to inferred conclusions. Unidirectionality is most apparent in deductive inference, but it is generally assumed to hold also for inductive inference. For example, Bayesian inference uses prior odds and likelihood ratios (premises) to derive posterior odds (conclusions). The direction of inference may vary depending on what is initially known (e.g., people may use known causes to infer potential effects or else use known effects to diagnose possible causes), but it is generally assumed that in any reasoning episode certain information constitutes the fixed premises from which certain other information can be derived as a (perhaps tentative) conclusion. The unidirectionality assumption rules out reverse inferences, that is, those that move from conclusions to premises.

Within this unidirectional framework, any inferences that seem to go against the flow are interpreted as signs of the frailty of human reasoning. In fact, apparent errors of this sort are prevalent. In deductive reasoning, people exhibit belief bias effects, allowing their prior beliefs about the truth or falsity of a proposed conclusion to influence their assessment of its validity (see Evans, 1989, for a review). When reasoning with conditionals, people often commit the fallacy of affirming the consequent, making a logically illicit inference from the conditional *If p then q* and the truth of the consequent, *q*, to the truth of the antecedent, *p* (e.g., Taplin & Staudenmayer, 1973). In the area of hypothesis testing, children—and to some extent adults as well—have difficulty distinguishing hypotheses to be tested from the evidence that should be used to evaluate them (Kuhn, 1989; Ranney, Schank, Hoadley, & Neff, 1996).

There is, however, an alternative conception of reasoning and decision making in which inferences are inherently bidirectional, so that the distinction between premises and conclusions is blurred. Bidirectional inferences are inherent in the operation of models of thinking that are based on parallel constraint satisfaction. Computational instantiations of such models are typically formulated as networks of units representing possibilities (e.g., possible beliefs or actions) that are interconnected by excitatory and inhibitory links representing positive and negative support relations between pairs of possibilities. Constraint satisfaction models operate by applying a relaxation algorithm, which settles the network into a stable state in which the asymptotic activation levels of the units define a set of winning possibilities (those with relatively high activation) that have succeeded in mutually supporting one another and collectively inhibiting their rivals. The bidirectional influences between related possibilities play a critical role in allowing the system to impose a coherent interpretation on an initially ambiguous set of inputs.

The interactive activation model of letter and word perception (McClelland & Rumelhart, 1981) was the prototype for computational constraint satisfaction models. This basic approach has been generalized to a variety of higher-

Keith J. Holyoak, Department of Psychology, University of California, Los Angeles; Dan Simon, Faculty of Law, Haifa University, Haifa, Israel.

We contributed equally to this project, and the order of authorship was determined alphabetically. This research was supported by National Science Foundation Grant SBR-95-11504. A preliminary report of this work was presented at the 37th Annual Meeting of the Psychonomic Society in Chicago, November 1996.

We thank Bobbie Naghdi Aguero, Xuan Bui, Eling Lin, and An Tran for assistance in preparing materials and testing participants and Jim Waltz for assistance with data analyses. Eric Melz programmed the extensions of the Explanatory Coherence by Harmany Optimization (ECHO) model used in the Coherence Model of Cognitive Consistency (Co3). Peter Bentler, Lien Pham, and Jodie Ullmann provided especially valuable assistance in conducting the structural equation modeling. Robert Branstrom, Michael Ranney, Tom Shultz, Bobbie Spellman, and Shelley Taylor provided helpful comments on an earlier draft of this article. The article also benefited from suggestions made by Harold Kelley and Paul Thagard.

Correspondence concerning this article should be addressed to Keith J. Holyoak, Department of Psychology, University of California, Los Angeles, Franz Hall, Los Angeles, California 90095-1563. Electronic mail may be sent to holyoak@lifesci.ucla.edu.

level cognitive processes, including analogical mapping (the Analogical Constraint Mapping Engine [ACME] model of Holyoak & Thagard, 1989), evaluation of competing explanations (the Explanatory Coherence by Harmany Optimization [ECHO] model of Thagard, 1989, 1992), and decision making (the Deliberative Coherence [DECO] model of Thagard & Millgram, 1995). At a more qualitative level, the constraint satisfaction approach has it roots in consistency theories developed in social psychology under the Gestalt influence. These consistency theories, which were applied to attitude and belief revision, included balance theory (Heider, 1946, 1958), dissonance theory (Festinger, 1957), and symbolic psycho-logic (Abelson & Rosenberg, 1958). The early consistency theories waned in influence over the years, in part because they were unable to specify how consistency could be reliably attained and because they were generally limited to networks that included just two or three elements. This problem has been solved by constraint satisfaction algorithms, which, for networks of any size, will adjust activations so as to reach an asymptotic state that maximizes the internal consistency (i.e., coherence) of the set of winning hypotheses. Classical consistency theories in social psychology have recently been resurrected and elaborated as computational models (Kunda & Thagard, 1996; Shultz & Lepper, 1996; Spellman, Ullman, & Holyoak, 1993; Read & Miller, 1994).

This article reports a series of experiments designed to explore the use of constraint satisfaction in reasoning about complex and ambiguous situations that require a decision about the most appropriate action to be taken, where the appropriateness of the action is determined by multiple inferences. In this type of task, people are required to make and integrate multiple inferences to reach a decision. Consider, for example, a person faced with a decision about whether to accept a job offer. Such a decision might be influenced not only by the starting salary, but also by an assessment of the causes of the company's recent growth, the reliability of the employer's assurances of promotion, the personality traits of prospective colleagues, a piece of professional advice regarding advantageous career changes, and the analogy of an acquaintance who recently took a similar job. In an interpersonal context, a person deciding how to respond to a nasty comment made by a friend might consider the possible motivation for the malevolent behavior, other circumstances that might have instigated the comment, the similarity between the comment and that friend's previous behavior, and the rules of appropriate conduct learned in kindergarten.

In everyday life, people routinely face such situations, which are laden with complexity, conflict, and ambiguity and yet seem to call for rationally justifiable inferences and decisions. In some situations of this sort, there may be no incontrovertible premises available; instead, the acceptability of each individual piece of evidence is intertwined with the acceptability of each other piece of evidence. When information about probabilities and utilities is nebulous, it is impossible to apply decision models that require precise quantitative information. Judgment heuristics, such as availability and representativeness (see Kahneman, Slovic, & Tversky, 1982), doubtless play a role; however, by themselves, such heuristics do not specify how conflicting information can be integrated.

The proposal we tested is that decisions that are based on complex but ambiguous information follow a function of maximal coherence among the underlying inferences and that the process of achieving coherence is based on constraint satisfaction, which depends on bidirectional links. The experiments we report involved college students acting as judges in a simulation of judicial reasoning. We wish to demonstrate (a) that coherence can, in fact, be achieved in the face of extreme ambiguity; (b) that the pressure to achieve coherence guides the decision-making process itself, rather than simply providing post hoc rationalizations (contrary to the view of Festinger, 1964); (c) that coherence-based shifts in beliefs and attitudes trigger correlated shifts in memory; and (d) that the impact of spreading coherence can extend through a chain of intermediate inferences to produce remote changes in beliefs.

Experiment 1

The main purpose of Experiment 1 was to address the first two questions raised above. To address the first question, we examined whether participants who were faced with a decision fraught with ambiguity shifted their beliefs so as to increase their coherence with one another and with the eventual decision. Participants were asked to evaluate a set of arguments first in isolation and later in the context of a legal case. In the latter instance, the inferences were phrased in the form of legal arguments, half of which were made by the plaintiff and half by the defendant. Hence, half of the arguments supported one decision, and the other half supported the opposite outcome. If decisions are based on bidirectional constraints between inferences, then we would expect assessments of the individual arguments to shift so as to lend stronger support for the final verdict. In addition, we would expect to observe a shift from zero or weak correlations among the argument evaluations in the preliminary assessments to robust positive correlations on the postdecision assessment.

A second condition tested in Experiment 1 addressed the second question by examining whether a shift in coherence precedes the generation of a verdict or only occurs after the verdict has been reached. If coherence among argument evaluations emerged prior to the decision, this finding would support the claim that development of internal consistency within a coherent position plays a causal role in reaching decisions, rather than arising as the result of postdecision efforts to rationalize and defend a verdict that had initially been reached for other reasons.

Method

Materials

The materials consisted of a legal case and sets of opposing arguments offered by the plaintiff and the defendant. The case was called *Caught in the Net* and involved a civil action triggered by a statement that was broadcast over the Internet. Although fictional,

this case was loosely based on an actual one described by Holyoak and Thagard (1995, p. 153). The dispute focused on a gray area of the law that had recently arisen out of the new technology of electronic communication over the Internet, which was not clearly regulated under existing laws. This situation made it possible to design a case that was fraught with ambiguity (including disputes over which legal precedent is most relevant) and was likely to be engagingly realistic to college students.

The case centered on a lawsuit launched by Quest, a software company, against Jack Smith, an investor in the company. Appendix B, Part II, gives the summary of the basic facts of the case that was read by participants. The facts, which were not in dispute, were that Quest's financial situation had deteriorated and its management was having difficulty in coping with the problems facing the company. Smith, a dissatisfied shareholder, posted a negative message about Quest's prospects on an electronic bulletin board directed at investors. Shortly thereafter, Quest's stock price plummeted and the company went bankrupt. It was later revealed that (unbeknownst to Smith) Quest had been secretly developing a new product that might have saved the company. Quest was now suing Smith for libel, claiming that his message caused the collapse of the company.

Each side made six arguments in favor of its position. The arguments formed opposing pairs, or points of dispute. The text of the arguments as presented to participants appears in Appendix B, Part III. The arguments for each side were parallel in form so as to encourage participants to align and compare the conflicting arguments for each point of dispute. (See Markman & Medin, 1995, for evidence that alignable arguments have greater impact on decisions than do arguments that are less clearly comparable.) The first three points of dispute involved matters of fact, and the second three involved matters of law or social policy. For expository convenience in this article, we have named each point of dispute, as follows. (Please note that the arguments were not explicitly labeled for participants.)

1. *Truth*: Quest argued that Smith's negative message was unfounded, whereas Smith claimed it was well-founded.

2. *Cause*: Quest asserted that the message caused the company's downfall, whereas Smith claimed that mismanagement was the cause.

3. *Motive*: Quest claimed that Smith's action was motivated by vindictiveness, whereas Smith claimed he only aimed to protect other innocent investors.

4. *Regulation:* Quest claimed that in posting his message, Smith had violated a company regulation requiring prior notification of management; Smith maintained that he had complied with the regulation.

5. *Speech*: Quest argued that it is in society's interest to regulate speech over the Internet, whereas Smith argued that society benefits from free speech over the Internet.

6. *Analogy*: Quest likened the Internet to a newspaper, which was subject to libel law, whereas Smith drew an analogy to a telephone system, which is immune from libel law.

The rival arguments for the case are presented verbatim in Appendix B, Part III.

We constructed two instruments to assess participants' opinions about each point of dispute. The first instrument (see Appendix A) was a pretest that was presented before participants were told about the Quest case. We constructed a total of eight questions, each presented with a brief context that was intended to correspond to the part of the Quest case relevant to that particular question. Each question was introduced as an independent query about "factual situations, public policy, business situations and legal affairs." Participants were told they were not expected to have any expert knowledge but were simply to use common sense in making their ratings. Each question's context introduced a distinct company or individual or else a general policy issue. For two of the points of dispute (Speech and Analogy), the opposing inferences were sufficiently independent that a separate question was constructed for each inference. For example, separate questions probed participants' assessment of the degree to which the Internet resembles a newspaper and to which it resembles a telephone system. For two other points (Cause and Motive), the opposing inferences were more closely related; these were probed by two-part questions. For example, after reading a description of the events leading up to an investor spreading a negative message, participants were asked to assess (a) whether he was motivated by vindictiveness and (b) whether he was motivated by a desire to protect other innocent investors. Finally, for two points of dispute (Truth and Regulation), the opposing inferences appeared to logically contradict one another. In these cases, only one question was asked. For example, after a company regulation and an investor's action were described, participants were asked to assess whether the investor had violated the regulation. All assessments were made by giving a rating on an 11-point scale, ranging from −5 (*strongly disagree*) to 5 (*strongly agree*), with a rating of 0 indicating neutrality.

The second instrument elicited participants' assessments of the parallel arguments in the context of the Quest case. These questions had the same form and wording as those used in the pretest except that they were now embedded in the Quest case and were formulated as arguments made by the two opposing parties. The wording of this instrument is presented in Appendix B, Part VI.

Design and Procedure

Participants were assigned to one of two conditions. Those in the 2-phase condition first completed the pretest. After this booklet was collected, participants spent 3–5 min completing an unrelated reasoning task. They then received a booklet that provided the factual summary and arguments for the Quest case. They were allowed to look back at the case as they went on to indicate a verdict and to provide a rating on a 5-point scale of their confidence that they had made the best possible verdict. Participants then completed the final posttest evaluation of the arguments.

Participants in the 3-phase condition completed the same procedure with the same materials as did those in the 2-phase condition, with the following additions. In the initial instructions for the Quest case (Appendix B, Part I), 3-phase participants were told that before reaching a verdict, they should wait to hear the verdict of another judge in a related case, as the other verdict would provide important additional information highly relevant to their decision in the case at hand. In the meantime, they were to read the case and think about it. The 3-phase participants were then asked to state their "preliminary leaning" toward either Quest or Smith, rating their confidence on a 5-point scale (Appendix B, Part IV). The second assessment instrument was then administered. After their response forms were collected, participants were then told that the other judge was not going to deliver a verdict after all (Appendix B, Part V) and that they should proceed to reach a final verdict based on the facts and arguments they had read. After stating their verdict, they completed the second assessment instrument again (with a different random order of questions).

Relative to those in the 2-phase condition, participants in the 3-phase condition thus provided an additional interim assessment of the points of dispute, after reading the Quest case but prior to being asked for a firm verdict. That is, the 2-phase condition involved a pretest and a posttest, whereas the 3-phase condition involved a pretest, an interim test, and a posttest. If a shift toward greater coherence was observed from the pretest to the interim test for the 3-phase participants, this would provide evidence that the

coherence shift preceded (and hence may have guided) the reaching of a decision, rather than simply following in its aftermath.

Within both the 2-phase and the 3-phase conditions, the order of the two sets of arguments (for Quest and for Smith) was counterbalanced across participants. Three different versions of each assessment instrument were created by forming different random orders of the eight questions (while always keeping the two parts of each two-part question together in a fixed order). The orderings were arranged so that for any participant, the corresponding questions appeared in a different ordering on each assessment instrument. Participants were tested alone or in groups of up to four. The entire experiment took from 30 to 50 min to complete.

Participants

Fifty-four undergraduates (35 women and 19 men) at the University of California, Los Angeles (UCLA), participated in the experiment in order to satisfy a course requirement. Twenty-four were assigned to the 2-phase condition and 30 to the 3-phase condition.

Results and Discussion

Preliminary inspection of the data revealed that six participants in the 3-phase condition gave a final verdict that differed from their preliminary leaning and that these "switchers" produced overall judgment patterns that differed from those obtained both for the other 3-phase participants and for the 2-phase participants. Hence, all the primary analyses were performed after excluding the 3-phase switchers, whose results are discussed separately. Thus, the primary analyses are based on data from 48 participants, 24 from the 2-phase condition and 24 from the 3-phase condition.

Distribution of Verdicts

The first concern was to establish whether the Quest case was open to alternative verdicts and whether participants were able to reach clear verdicts despite the inherent ambiguity created by the conflicting arguments. The distribution of verdicts did not differ significantly between the 2-phase and 3-phase conditions; accordingly, we report the aggregate results collapsing across conditions. Participants were about evenly divided in their verdicts, with 26 deciding in favor of the plaintiff, Quest, and 22 deciding in favor of the defendant, Smith. Yet despite the apparent ambiguity of the case as evidenced by the even split of verdicts, individual participants were generally very confident that they had reached the best possible decision. Figure 1 displays the distribution of confidence ratings for the 48 participants. Seventy-five percent of participants indicated that they had maximal (5) or next-to-maximal (4) confidence in their verdicts; conversely, only 5% indicated they had minimal (1) or next-to-minimal (2) confidence. This combination of ambiguity (evenly divided verdicts) and high individual confidence in decisions is consistent with constraint satisfaction models of decision making, according to which ambiguous situations are resolved by allowing one coherent set of beliefs to become highly activated, inhibiting the rival set

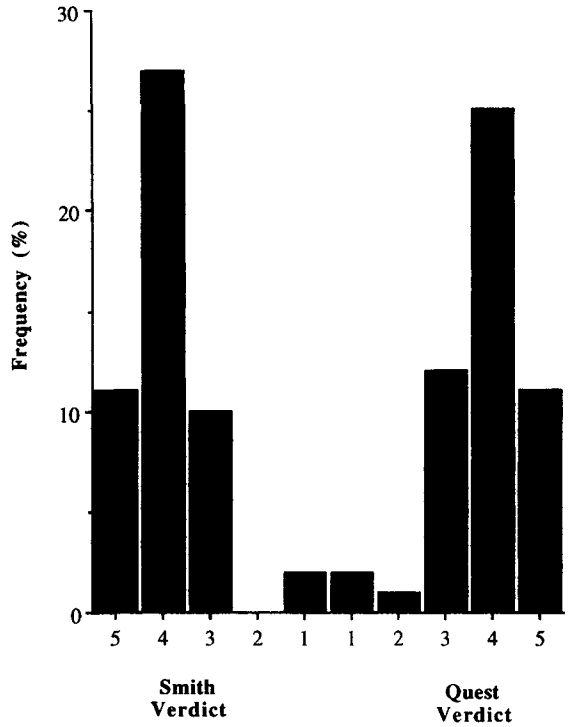


Figure 1. Distribution of confidence ratings for verdict in Experiment 1. Ratings are on a 5-point scale, ranging from 1 (*minimal confidence*) to 5 (*maximal confidence*).

(much like the two alternative perceptual interpretations of the well-known Necker cube).

Shifts in Assessment of Points of Dispute

The next question to be addressed concerned whether the process of reaching a verdict was accompanied by shifts in participants' assessments of the six points of dispute between the plaintiff and the defendant. Constraint satisfaction models of decision making predict that an emerging decision will be accompanied by a general shift toward a coherent position across all the points of dispute.

In order to measure participants' positions on each disputed point, we converted the ratings obtained for each assessment instrument (pretest for both conditions, interim test for 3-phase condition only, and posttest for both conditions) to values we termed *Q-scores*, which provide an index of agreement with the position of the plaintiff, Quest. We computed the Q-score for each point of dispute by taking a weighted average of the ratings for questions that assessed that point, reversing the scale for those questions for which positive values indicated support for Smith's position (see Appendix A and Appendix B, Part VI). For example, the Q-score for Truth was simply the negation of the rating for the question of whether Smith's message was well-founded (e.g., a rating of 5, indicating high agreement with Smith's claim that the message was true, was converted to a Q-score of −5, indicating minimal support for Quest's position on this issue). In a similar manner, the Q-score for Analogy was

the mean of the rating in support of the newspaper analogy (Quest's position) and the negation of the rating in support of the telephone analogy (Smith's position). All Q-scores therefore ranged from −5 (minimal support for Quest's position) to 5 (maximal support for Quest's position), with 0 indicating neutrality. Mean Q-scores were calculated by averaging the Q-scores for the individual points of dispute.

Figure 2 presents the mean Q-score on each assessment, plotted separately for participants who decided in favor of Quest versus those who decided in favor of Smith. It is clear from inspection of Figure 2 that the two groups had similar Q-scores on the pretest. The distribution of these pretest scores was approximately normal, with 85% of participants having a mean between −1 and 1. However, the Q-scores for the two groups sharply diverged in the direction of the verdict on the interim test (3-phase condition only) and on the posttest (both conditions). This apparent pattern was supported by the results of analyses of variance (ANOVAs). We performed two sets of analyses. The first set of analyses considered the comparable data for the 2-phase and 3-phase conditions (i.e., the pretest and the posttest). The shift in the Q-score across tests, in opposite directions for the Quest versus Smith participants, yielded a highly significant interaction, $F(1, 44) = 35.3$, $MSE = 22.6$, $p < .001$. Tests of simple main effects revealed that the difference in Q-scores on the pretest between participants who would eventually decide in favor of Quest versus Smith fell short of significance, $p = .06$, although a trend was apparent (mean Q-score of .37 for Quest supporters versus −.20 for Smith supporters). In contrast, the difference between the two groups was highly significant on the posttest (*M*s = 1.42 and −1.56, respectively), $p < .001$. The 2-phase and 3-phase conditions did not differ overall, nor did they differ with respect to the shift in Q-scores across tests, $F < 1$ for both tests. The shift in Q-scores as a function of verdict also did not vary significantly across the six points of dispute, $F(5, 22) = 1.53$, $MSE = 11.6$, $p = .18$. Q-scores for each of the six individual points shifted in the direction that cohered with the verdict.

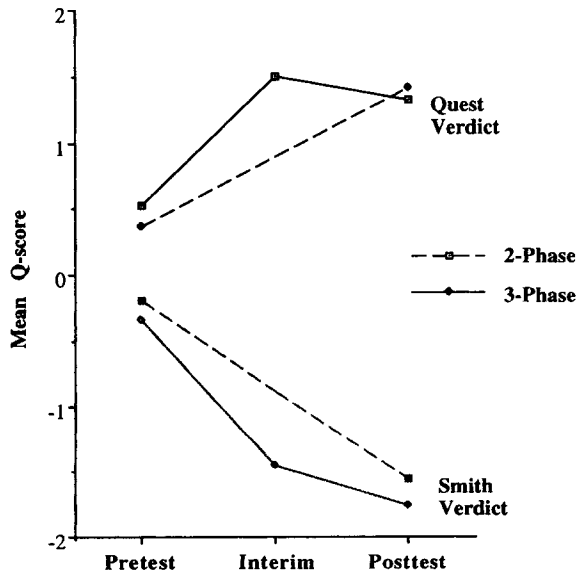


Figure 2. Shifts in Q-scores (favorability to Quest's position) across tests as a function of eventual verdict for Quest versus Smith in Experiment 1. Lower scores indicate less favorability to Quest's position, whereas higher scores indicate more favorability to Quest's position.

We performed a second ANOVA for the data from the 3-phase condition only, assessing the change across the three tests (pretest, interim test, posttest). The interaction between verdict and change in mean Q-scores across tests was also highly significant in this analysis, $F(2, 44) = 11.2$, $MSE = 11.2$, $p < .001$; again, this interaction did not vary across the six points of dispute, $F < 1$. Newman-Keuls tests revealed that the shifts from the pretest to the later tests were all highly significant, $p < .01$, but that the Q-scores on the interim test and the posttest did not differ, $p > .20$. These results thus indicated that the major shift in participants' assessments of the points of dispute occurred before they presumably reached a verdict, thus supporting the hypothesis that the emergence of coherence occurs in the process of reaching a decision, rather than in its aftermath.

Although Q-scores in the 3-phase condition did not increase from the interim test to the posttest, a separate analysis revealed a modest increase in verdict confidence from the preliminary leaning on the interim test ($M = 3.36$ on the 5-point scale) to the final verdict on the posttest ($M = 3.91$), $t(21) = 5.02$, $p < .001$. Final verdict confidence did not differ between the 3-phase and the 2-phase conditions (3.91 versus 3.86), $t < 1$.

Even though the pretest Q-scores did not differ significantly between the eventual Quest and Smith supporters, the trend apparent in Figure 2 suggested that it would be worthwhile to examine the data in more detail to see if any individual pretest questions appeared to predict the eventual verdict rendered. Thus, we conducted an ANOVA on the ratings for the 10 pretest questions, and we tested the difference between the two verdict groups for each question. Only one pretest question proved to be significantly linked to the verdict. This was one of the questions used to define the Speech factor, namely, the assessment of the proposition, "As a matter of policy, communications over the Internet ought to be regulated by law." Participants who would later decide for Quest were significantly more favorable to this statement than were those who would decide for Smith (*M*s = 1.96 and −1.09, respectively), $F(1, 46) = 17.21$, $MSE = 6.45$, $p < .001$. It was interesting that the two subgroups did not differ in their ratings of the other Speech statement, "As a matter of policy, keeping the Internet open to the free exchange of viewpoints and information is a vital social need" (*M*s = 2.38 and 2.55, respectively), $F < 1$. It seems that people in both subgroups were willing to endorse the general idea of keeping the Internet free, but those who would decide for Quest were more willing to also express support for the idea that the Internet should be regulated. This pretest difference on the issue of regulating the Internet might have been the seed (at least for some participants) that initiated the eventual opinion shifts accompanying the process of reaching a verdict. As this particular pretest

difference was not predicted, we sought replications (generally successfully) in subsequent experiments.

Shifts in Coherence Among Points of Dispute and Verdict

Although the above analyses revealed a clear shift in participants' assessments of the six points of dispute in the direction of their verdict, they do not suffice to establish that individual participants reached a broadly coherent position across the disputed points. It remains possible, for example, that each participant was eventually persuaded by some single argument for one side in the case, with the particular critical argument varying from one person to the next. However, if a constraint satisfaction process was used to reach a decision, then individual participants would be expected to shift their assessments of most or all of the disputed points in the direction of their eventual verdict.

Such a general increase in coherence could be revealed by a correlational analysis. On the pretest, participants' assessments of the six positions would not constrain one another and, hence, would tend to be uncorrelated. Once the points are presented in the context of the case, however, a constraint network would be created, the effect of which would be to generate positive correlations among the disputed points and between each point and the verdict.

The results of such a correlational analysis are displayed in Table 1. These correlations are based on the pretest and posttest Q-scores for the combined data from the 2-phase and 3-phase conditions (as the patterns for the two conditions did not differ in any major way). The verdict was treated as a binary variable (1 for a Quest verdict, 0 for a Smith verdict). (Extremely similar results were obtained when the verdict was treated as a continuous variable based on verdict confidence.) The increase in coherence was striking. On the pretest, only 2 of the 21 correlations among the disputed points and verdict were significantly positive, and several were negative. (Note that one of the positive correlations was between pretest Speech and the verdict, reflecting the predictiveness of the "regulate Internet" question.) In contrast, on the posttest, all but one of the correlations were significantly positive, including all six correlations between disputed points and the verdict; the nonsignificant correlation was also positive. A similar analysis for the 3-phase condition alone revealed that the increase in correlations, like the shift in Q-scores, occurred from the pretest to the interim test, with only small additional increases in the correlations between the interim test and the posttest. The correlational analyses thus provided further evidence that the emergence of a coherent position on the disputed points led, rather than followed, participants' arrival at a firm decision.

In order to provide a visual depiction of the coherence increase apparent in the correlation matrices in Table 1, we submitted a half-matrix consisting of the 78 correlations among 13 factors (six pretest Q-scores, six posttest Q-scores, and verdict) to ALSCAL (Young & Harris, 1990), a computer program that performs nonmetric multidimensional scaling (Shepard, 1962). We transformed the correlations into ordinal measures of dissimilarity by subtracting each from 1. Figure 3 presents a plot of the ALSCAL solution (with a stress value of .17, using Kruskal's Formula 1) placing the 13 factors in a two-dimensional similarity space. We did not attempt to interpret the axes (which could be freely rotated) but, rather, examined the emergence of a cluster of factors on the posttest. To highlight the emergence of coherence in the posttest, the six posttest positions in Figure 3 are enclosed within a convex border (estimated by eye). Whereas the pretest positions are scattered around the

Table 1
Correlations Among Final Verdict and Positions on the Six Points of Dispute: Experiment 1

Point of dispute	1	2	3	4	5	6	7
				Pretest			
1. Truth	—	−.02	.35*	.11	.19	.10	.09
2. Cause		—	−.06	.28	−.09	−.30*	.14
3. Motive			—	.07	−.14	.03	.05
4. Regulation				—	.18	−.05	.11
5. Speech					—	.24	.42**
6. Analogy						—	−.03
7. Verdict							—
				Posttest			
1. Truth	—	.34*	.50**	.31*	.40**	.39**	.54**
2. Cause		—	.55**	.45**	.34*	.53**	.68**
3. Motive			—	.39**	.24	.47**	.52**
4. Regulation				—	.43**	.43**	.69**
5. Speech					—	.57**	.63**
6. Analogy						—	.68**
7. Verdict							—

*$p < .05$. **$p < .01$.

plot (with pretest Speech falling closest to the verdict), the posttest positions form a compact cluster. Furthermore, the verdict falls extremely close to the center of the posttest cluster. This visual representation captures the central properties of the correlational data: Initially unrelated positions on the disputed points were transformed over the course of decision making into a coherent position closely allied with the final verdict.

Coherence shifts were also observed *within* one of the theoretically most interesting factors, Analogy. Figure 4 presents the shifts in support ratings for the questions assessing the degree to which the Internet is like a newspaper (a pro-Quest position) and like a telephone system (a pro-Smith position), plotted separately for participants who decided in favor of Quest versus Smith (collapsing over the 2-phase and 3-phase conditions, which did not differ reliably). Regardless of their eventual verdict, participants on the pretest gave very similar and mildly favorable ratings (ranging from .82 to 1.12) to both analogies, and the ratings tended to have a weak positive correlation, $r(46) = .28, p < .10$. By the posttest, however, the assessments strongly diverged in the direction cohering with the verdict: Quest supporters now were much more favorable to the newspaper analogy and negative to the telephone analogy, whereas Smith supporters offered the reverse assessments. An ANOVA revealed that this three-way interaction among the factors Verdict, Test, and Analogy was highly significant, $F(1, 46) = 19.6, MSE = 146, p < .001$. The correlation between the assessments of the two analogies on the posttest was strongly negative, $r(46) = -.65, p < .01$. Thus, as participants worked their way to a verdict, the two analogies that were initially evaluated quite independently came to be viewed as rivals. One analogy eventually prevailed along with the verdict with which it cohered, and the rival analogy came to be viewed negatively. This shift in preferences between alternative possible analogies created a possible vehicle for long-distance transfer effects in the aftermath of reaching a decision that determined the preferred analogy. We investigated this possibility in Experiment 3.

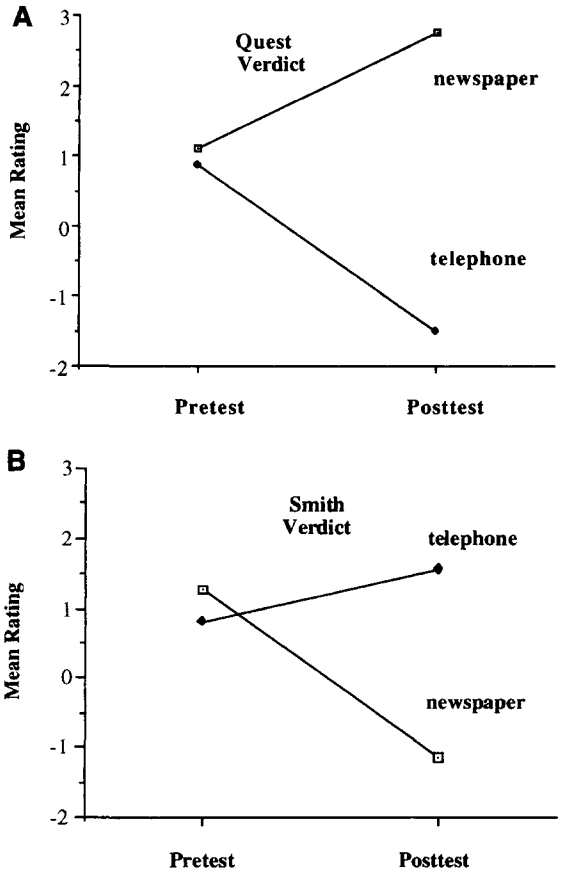


Figure 4. Shifts in ratings of the newspaper and telephone analogies across tests as a function of eventual verdict for Quest (A) versus Smith (B). Ratings were on an 11-point scale, ranging from −5 (*strongly disagree*) to 5 (*strongly agree*).

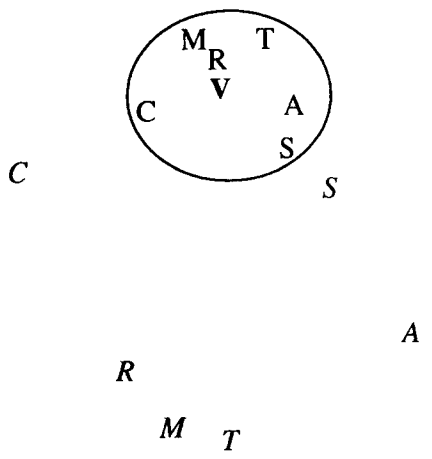


Figure 3. Two-dimensional nonmetric scaling solution for correlation matrix among positions on the disputed points in Experiment 1, displaying the shift from unrelated initial positions on the pretest (italicized) to a coherent position associated with the verdict on the posttest (Roman type). V = verdict; A = analogy; C = cause; M = motive; R = regulation; S = speech; T = truth.

Although the Analogy factor produced the largest within-factor coherence shift, similar shifts were observed for other factors. Furthermore, the negative shifts tended to be larger for within-factor comparisons (e.g., regulate-Internet vs. free-speech assessments) than for between-factor comparisons (e.g., regulate-Internet vs. telephone-analogy assessments). We examined the pretest and posttest correlations between assessments of the two questions for the Cause, Motive, Speech, and Analogy factors (the factors assessed with two questions), calculated separately for within-factor and between-factor comparisons. The mean correlations for within- and between-factor comparisons were −.09 and −.04, respectively, on the pretest and −.46 and −.24, respectively, on the posttest. The fact that the posttest correlations were more strongly negative for within-factor than between-factor comparisons indicated that the directly opposing arguments tended to be viewed as more incompatible with one another than were arguments that only opposed each other indirectly by favoring opposite verdicts.

Three-Phase Switchers

Because only 6 of 30 participants in the 3-phase condition switched their verdict between the interim test and the posttest, the data for this small group of so-called switchers are too unreliable to merit being reported in detail. However, the qualitative aspects of their performance are of some interest. The very fact that a small but substantial minority (20%) of participants changed their minds between the interim test and the posttest indicates that the preliminary leanings on the interim test were not invariably firm decisions. Hence, it seems reasonable to conclude that we were successful in our efforts to encourage participants in the interim phase of the 3-phase condition to think carefully about the case without committing themselves to a verdict. As no new information relevant to their decision was ultimately presented between the interim test and the posttest, it is not surprising that the majority of participants eventually reached a final verdict consistent with their preliminary leaning. The fact that participants' assessments of the disputed points shifted between the pretest and the interim test can be interpreted as evidence that the coherence shift is inherent to the decision-making process itself, rather than only being invoked after participants have committed themselves to a firm decision.

Of the six switchers, three moved from an initial leaning toward Quest to a final verdict for Smith, and three moved in the opposite direction. The three who moved from Smith to Quest tended to maintain coherence in their assessments of the disputed points (favoring Smith's position on the interim test but Quest's position on the posttest). However, the three who switched from Quest to Smith failed to maintain coherence. In fact, the final Q-scores of the two subgroups (those who finally decided for Quest versus Smith) were equal (2.06). An important issue for future research will be to investigate the causes and consequences of such decision reversals.

Experiment 2

The coherence shifts observed in Experiment 1 occurred after a very brief (3–5 min) delay between the pretest and presentation of the Quest case. Experiment 2 was a small-scale replication of the 2-phase condition of Experiment 1, using a considerably longer delay (1 week). In addition, Experiment 2 examined whether changes in assessments of arguments, triggered during the process of reaching a verdict, would retroactively influence participants' memory for their earlier assessments given on the pretest. Research on memory for attitudes indicates that attitudinal change often produces a bias to recall one's earlier attitude as being closer to one's current attitude than was actually the case (Bem & McConnell, 1970; Goethals & Reckman, 1973; Ross & Shulman, 1973; Wixon & Laird, 1976). It follows that one consequence of coherence-based decision making is that participants' memory for their entering positions about disputed points will be shifted toward their final positions, which cohere with the verdict.

Method

Participants

Nineteen UCLA undergraduates (15 women and 4 men) who were enrolled in a upper-division laboratory course in cognitive psychology participated in the experiment during three class periods, each 1 week apart. All participants were volunteers, and the study was conducted by one of their classmates.

Materials, Design, and Procedure

The basic materials were virtually identical to those used in the 2-phase condition of Experiment 1. Because of time constraints, the Quest case was shortened by eliminating reference to a company regulation. Questions related to the regulation were removed from all instruments, leaving five points of dispute.

During Session 1, participants completed the pretest instrument. One week later, during Session 2, they read the Quest case, gave their verdict, and completed the posttest instrument. An additional week later, during Session 3, participants were given a new version of the pretest (with the questions in a different random order). They were asked to write down the same ratings they had given on the initial pretest. To be specific, the instructions stated, "For each question, your goal is to state the rating that you gave on the earlier test. Note that you should *NOT* give the rating you might now believe is correct (since your opinions might have changed). Rather, you should try your best to remember what rating you gave previously, and give that same rating again." Participants had not been told of this final recall task in advance. Sessions 1 and 3 lasted approximately 10 min, and Session 2 lasted approximately 20 min.

Results and Discussion

The 19 participants split evenly in their verdicts, with 9 deciding in favor of Quest and 10 in favor of Smith. All but one participant (95%) gave a rating of verdict confidence in the medium to high range (3–5 on the 5-point scale). The basic coherence shifts observed within a single session in Experiment 1 were replicated between Sessions 1 and 2 in Experiment 2, with a 1-week delay. Figure 5 displays the shift in Q-scores from pretest to posttest for participants who decided in favor of Quest. The pretest Q-scores were virtually identical for the two subgroups (−.34 for Quest supporters, −.30 for Smith supporters), whereas the posttest scores sharply diverged so as to cohere with the verdict (.94 for Quest supporters, −1.03 for Smith supporters). This interaction between verdict and test was reliable, $F(1, 17) = 7.71$, $MSE = 6.26$, $p = .01$, and did not vary significantly across the five disputed points, $F < 1$. We performed a separate analysis to find if any of the individual pretest questions predicted the eventual verdict. As in Experiment 1, participants who eventually decided for Quest tended to give a higher rating of support for regulating the Internet than did participants who eventually decided for Smith (*M*s = 1.00 and −.20, respectively). However, neither this difference nor any other difference between pretest questions was reliable, perhaps because of the smaller sample size in Experiment 2.

Table 2 presents the correlational analysis of the coherence shift, performed in the same manner as in Experiment 1. Although the shift was less statistically robust than in

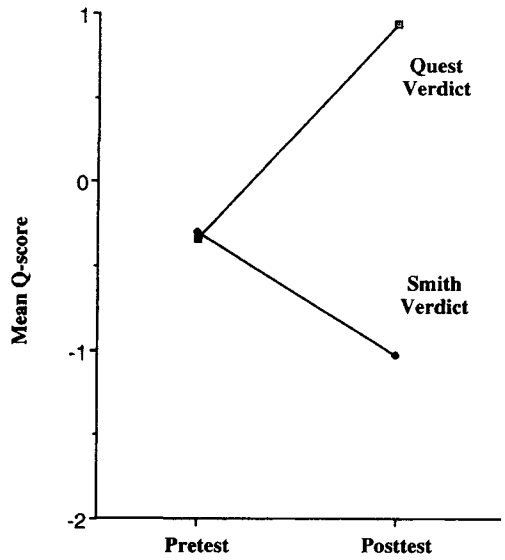


Figure 5. Shifts in Q-scores (favorability to Quest's position) across tests as a function of eventual verdict for Quest versus Smith in Experiment 2. Lower scores indicate less favorability to Quest's position, whereas higher scores indicate more favorability to Quest's position.

Experiment 1 (likely attributable to the smaller number of participants in Experiment 2), the magnitudes of the correlations were very similar. On the pretest, none of the correlations among the positions on points of dispute had a significant positive correlation with one another or with the eventual verdict. About half of the trends were negative, with one significant negative correlation. In contrast, all of the correlations on the posttest were positive, and several (mainly involving Analogy) achieved statistical significance.

Memory for pretest ratings was assessed by computing Q-scores on the basis of the recall test (administered in Session 3, after a 2-week delay and the intervening decision task performed in Session 2). One participant who failed to complete the memory test was excluded from these analyses. The mean recalled Q-score was −.36 for the Quest supporters, very close to their actual mean pretest Q-score of −.34; however, the mean recalled score for the Smith supporters, −1.32, was lower than their actual mean score of −.30. The difference between the recalled Q-scores for the two subgroups approached significance, $F(1, 16) = 3.66$, $MSE = 1.12$, $p = .07$. Across individual participants, the best predictor of the recalled Q-score was actually the verdict, $r(16) = .43$, $p = .07$, rather than the actual pretest Q-score, $r(16) = -.20$, $p > .4$, or the posttest Q-score, $r(16) = .26$, $p > .2$. Thus, although the small number of participants limited the reliability of the results, the data were consistent with the hypothesis that recall of initial positions is biased by the intervening decision. Experiment 3 further explored the impact of decision making on memory for initial positions.

Experiment 3

Experiment 3 had three major, interrelated objectives. First, it sought to experimentally manipulate the process of coherence-based decision making so as to examine the remote influences of variations in the input to the decision mechanism. In the previous experiments, all participants saw a single highly ambiguous case and reached a verdict; on the basis of that verdict, we then divided participants into subgroups and examined correlated shifts in assessments of the points of dispute. In contrast, participants in Experiment 3 were presented with one of two variants of the Quest case, each designed to experimentally induce a different verdict. These variants generally presented the same facts; however, they differed with respect to information about the personal history and character of the defendant, Jack Smith. In general, in the so-called Good Smith version the defendant

Table 2
Correlations Among Verdict and Positions on the Five Points of Dispute: Experiment 2

Point of dispute	1	2	3	4	5	6
			Pretest			
1. Truth	—	−.02	−.19	.07	.01	−.26
2. Cause		—	−.52*	−.07	−.41	−.11
3. Motive			—	.25	.21	.22
4. Speech				—	−.08	.12
5. Analogy					—	.02
6. Verdict						—
			Posttest			
1. Truth	—	.05	.45	.18	.31	.55*
2. Cause		—	.34	.31	.60**	.40
3. Motive			—	.38	.58**	.43
4. Speech				—	.60**	.39
5. Analogy					—	.63**
6. Verdict						—

*$p < .05$. **$p < .01$.

had a history of honest concern with the operation of companies in which he invested, whereas in the so-called Bad Smith version he had a history of unscrupulous manipulation. These differences were intended to bear directly on one of the points of dispute, Motive. That is, Good Smith would appear likely to have had an altruistic motive for posting the negative message, whereas Bad Smith would appear likely to have acted out of greed, malice, or vindictiveness. If the experimental manipulation was effective, it would have an impact on participants' assessment of the Motive factor and possibly also on their verdict. These effects would be predicted by any plausible account of decision making and would simply serve as a manipulation check.

However, if decisions are truly coherence based, the impact of the manipulation would be considerably more extensive. Much as forcing one interpretation on a vertex of a Necker cube can cause the entire set of vertices to be perceived as a particular three-dimensional form, biasing the assessment of one particular point of dispute should initiate spreading coherence, causing systematic changes in the entire set of assessments related to the verdict. In a constraint satisfaction network, if one argument unit (e.g., Quest's argument that Smith's motive was malevolent) becomes highly active, it will inhibit the rival argument with which it directly competes and excite the verdict with which it coheres (e.g., a decision for Quest). The favored verdict unit will, in turn, inhibit the opposing verdict while exciting all the other argument units with which it coheres (e.g., Quest's claims that Smith's message caused its downfall and that the Internet should be likened to a newspaper). Because the links between a verdict and the arguments with which it coheres are bidirectional, synergistic feedback loops will operate, affording mutual support to both the favored verdict and its supporting arguments. The network will eventually settle into an asymptotic state in which one coherent position is highly active (e.g., strong support for a verdict for Quest coupled with positive assessments of all its supporting arguments) and the rival position is suppressed (e.g., weak support for a verdict for Smith coupled with negative assessments of all its supporting arguments). If such spreading coherence effects were in fact obtained, they would establish that coherence-based decision making can be alogical, yielding inferences that would be difficult to explain by means of any logical calculus. For example, an inference from the shady history of the defendant to the conclusion that the Internet resembles a newspaper more than it does a telephone system is coherent in this particular constellation but is not logically compelling.

The second major objective of Experiment 3 was to determine whether spreading coherence could bias inferences and decisions in a subsequent transfer case. We have argued that in the course of reaching a decision by constraint satisfaction, people will change their assessments of the points of dispute. People's final assessments could then influence the constraint network for a subsequent problem in which some of the same arguments are again relevant. Analogy is a particularly plausible candidate for generating such transfer effects. In Thagard's (1989, 1992) ECHO model of evaluating competing explanations, for example, analogy is one source of constraints. The nature of analogical reasoning is to use knowledge of the source analog to generate parallel inferences about the target. If the source analog is a rich one, there are an indefinite number of plausible analogical inferences that might be generated if they are contextually relevant to a new target problem.

One of the key points of dispute in the Quest case concerned whether the better source analog for the Internet is a newspaper or a telephone system. Suppose, following the example used previously, that a person decides for Quest and in doing so comes to believe that the Internet is basically an electronic newspaper. If a different case is then presented—even one that has little overlap with the issues involved in the Quest case—in which the choice of source analog for the Internet is relevant, the person may be predisposed to again favor the newspaper source analog over the telephone alternative. If so, the favored analogy may provide a bridge that allows coherence to spread from the Quest case to the transfer case, thereby triggering additional inferences and decisions in the latter case that will tend to cohere with the person's final position on the Quest case.

The third objective of Experiment 3 was to explore further whether coherence-based decision making biases people's later recall of their entering positions on the points of dispute. Participants completed the entire study in a single session, with about a 30-min delay between the pretest and a recall test of memory for the pretest ratings (as compared with the 1-week delay used in Experiment 2). If recalled pretest ratings were biased by the intervening decision process even when the delay was relatively brief, this finding would provide stronger evidence that people's memories for their initial positions shift so as to cohere with their new positions.

Method

Materials

The materials for the Quest case were based on those used in Experiment 1 (including the point of dispute regarding whether or not Smith violated the company regulation, which was omitted in Experiment 2). The only change was in the factual summary, for which two new variants were written, the Good Smith and Bad Smith versions. (These names are simply for ease of exposition; they were never presented to participants.) As sketched above (and reproduced verbatim in Appendix C, Part I), Good Smith had a history of constructive criticism of companies in which he invested, whereas Bad Smith (Appendix C, Part II) had a history of harming and libeling them.

In addition, materials for a second case, *The Bonus Dispute at Infoscience*, were written to serve as a transfer problem. These materials appear in Appendix D. This case involved a contract dispute between a company that runs a bulletin board on the Internet and its employees with regard to how high a bonus should be paid to employees. The contract specified that the bonus should be related to two factors: the bonus paid at similar information service firms located in the vicinity and the extent to which the company's profits could be attributed to the employees' efforts. The arguments by the two sides focused on these two points of dispute. The first factor, Analogy, concerned whether the most similar

company to Infoscience was the local newspaper or the local telephone company. Which side cited which analogy was counterbalanced (with the analogy cited by Infoscience always supporting a lower bonus than that cited by the employees). The two sides gave legalistic definitions of newspaper and telephone system that were identical to those used in the Quest case. Thus, the Analogy factor was a shared point of dispute that served as a bridge between the Quest and Infoscience cases.

The second point of dispute, Credit, was specific to the Infoscience case. This point concerned whether the company's profits were mainly attributable to installation of a new computer system (the position of Infoscience) or to the efforts of the employees (the position of the employees). The Infoscience materials had the same overall structure as the Quest materials: opening instructions, factual summary, opposing arguments, elicitation of a verdict, and a posttest concerning its two points of dispute. In addition, two questions related to Credit were added to the initial pretest (which preceded both cases).

Design and Procedure

The general procedure was extremely similar to the 2-phase condition of Experiments 1 and 2. Participants were assigned to one of three treatment groups. The two main conditions differed solely in whether the factual summary for the Quest case was the Good Smith or the Bad Smith version. Participants in both these groups performed the following tasks in order: (a) pretest, (b) a 3–5 min unrelated reasoning problem as a filler task, (c) Quest case (reading summary and arguments, giving verdict, and completing posttest), (d) Infoscience case (with same structure as the Quest case), and (e) recall test of ratings given on the pretest.

A third group performed only tasks (a), (b), (d), and (e) above (i.e., the Quest case was omitted). This group was included to provide a baseline for verdict distributions on the Infoscience case in the absence of potential transfer from the Infoscience case.

As in the previous experiments, questions on the pretest, posttest, and recall test were presented in a distinct random order, and the assignment of question order was counterbalanced across participants. In addition, the order of the plaintiff's and defendant's arguments was counterbalanced for both the Quest and the Infoscience cases. Whereas in the Quest case the company always cited the newspaper analogy, in the Infoscience case which side cited each analogy was counterbalanced. Thus, any transfer between the two cases based on the bridging analogy could not be attributed to a general bias for or against corporations (e.g., a verdict for Quest in the first case, which should lead to transfer based on the newspaper analogy, would for half the participants cohere with a verdict *against* Infoscience, which in their version would have cited the telephone analogy).

Participants were tested alone or in groups of two to five people. The entire study required approximately 50 min for those in the two experimental conditions and about 30 min for those in the baseline condition.

Participants

Participants were 80 UCLA undergraduates who completed the study to satisfy a course requirement for a psychology class. Thirty-two participants were assigned to each of the two experimental conditions, and 16 were assigned to the Infoscience baseline condition.

Results and Discussion

We first report data and analyses concerning the verdict and points of dispute in the Quest case. Then, we report the results concerning coherence-based transfer from the Quest to the Infoscience case as well as the results of fitting structural equation models to the pattern of verdicts and assessments across both cases. Finally, we report the results concerning recall memory for pretest ratings. All of these analyses focus on coherence effects.

Paths of Spreading Coherence Within Quest Case

Direct influences of the character manipulation. For the Quest case, we first examined whether the manipulation of the defendant's character influenced the posttest assessment of the Motive factor, the point of dispute to which the character information was most directly relevant. As expected, participants who read the Bad Smith version were much more favorable to Quest's claim that Smith was malevolent (mean Q-score of 1.83) than were participants who read the Good Smith version ($M = -2.75$), $t(62) = 8.98$, $p < .001$. The experimental manipulation also had a strong impact on the verdict for the Quest case, with 78% of participants in the Bad Smith condition deciding in favor of Quest versus 28% in the Good Smith condition, $\chi^2(1, N = 64) = 16.1$, $p < .001$. These results thus confirmed that the experimental manipulation was effective in influencing participants' verdicts.

Influences of the character manipulation on assessments of disputed points. We then examined the influence of the character manipulation on the remaining points of dispute, to which it was linked less directly. Figure 6 presents the mean changes in Q-scores from pretest to posttest for these points. A clear interaction is apparent. On the pretest, the mean Q-scores were very similar for the Bad Smith and Good Smith conditions (−.27 versus .01), whereas the posttest Q-scores sharply diverged to cohere with the character manipulation (.93 for Bad Smith versus −.71 for Good Smith). We performed two sets of analyses to confirm this apparent interaction. Because of a clerical error, the pretest question concerning the Regulation factor was omitted for a third of the experimental participants. Hence, all analyses involving this factor were limited to the 36 participants who received this pretest question. An ANOVA including only the Regulation factor yielded a highly significant interaction between condition and test, $F(1, 34) = 9.76$, $MSE = 6.36$, $p < .01$. Q-scores for the two conditions did not differ significantly on the pretest but were higher for the Bad Smith than the Good Smith condition on the posttest, $p < .001$. A comparable analysis for the remaining points of dispute (the Truth, Cause, Speech, and Analogy factors) also yielded a significant interaction, $F(1, 62) = 16.0$, $MSE = 5.19$, $p < .001$. This interaction had the same general form for all four points, but the size of the coherence shifts varied across the points, $F(3, 186) = 4.14$, $p < .01$. Test of simple main effects revealed that on the pretest, the two experimental conditions did not differ significantly for any of the points. On the posttest, Q-scores

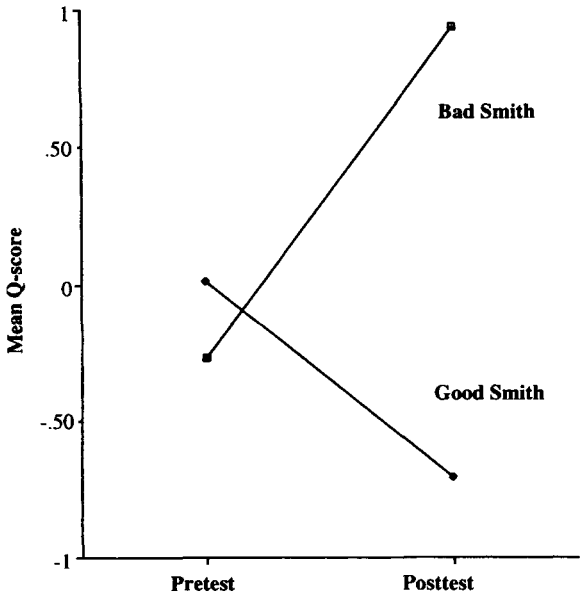


Figure 6. Shifts in Q-scores (favorability to Quest's position) across tests as a function of defendant's character in Experiment 3. Lower scores indicate less favorability to Quest's position, whereas higher scores indicate more favorability to Quest's position.

were significantly higher ($p < .01$) in the Bad Smith than the Good Smith conditions for Truth and Cause but fell short of significance for Analogy ($p = .12$) and Speech ($p = .28$). Overall, the results supported the prediction that the impact of the differential character information would extend to points of dispute to which it is only indirectly linked by means of coherence relations.

Shifts in coherence among points of dispute and verdict for Quest case. Table 3 presents the correlations among the Quest verdict and the points of dispute (leaving aside Motive, for which the pretest and posttest questions were not comparable because of the experimental manipulation). The results closely replicated those obtained in Experiments 1 and 2. On the pretest, only 2 of the 15 correlations were significantly positive (including that between pretest Speech and the verdict), whereas on the posttest all were. Figure 7 presents the corresponding two-dimensional nonmetric scaling solution (stress = .14), which again shows the shift from scattered pretest positions on the disputed points to a compact posttest cluster that includes the verdict. We performed a separate analysis to assess whether any of the pretest questions individually predicted the eventual verdict. As in the previous experiments, eventual Quest supporters were more favorable to regulation of the Internet than were eventual Smith supporters (*M*s = 2.26 and .40, respectively), $F(1, 62) = 6.69$, $MSE = 8.29$, $p = .01$. No other pretest question was a reliable predictor of the Quest verdict. Note that pretest Speech (the factor based in part on the "regulate Internet" question) lies relatively close to the verdict in Figure 7.

We also performed a correlational analysis to assess the predicted shift in the relationship between assessments of the newspaper and telephone analogies. On the pretest, the ratings of the two analogies were statistically independent, $r(62) = -.14$, $p = .28$. On the posttest, however, assessments of the two analogies had a significant inverse relationship, $r(62) = -.42$, $p < .01$.

Paths of Spreading Coherence Connecting Quest Case to Infoscience Case

Shifts in coherence among points of dispute and verdict for Infoscience case. We performed a similar series of analyses to examine shifts in internal coherence among the verdict in the transfer case, Infoscience, and its two points of dispute, Analogy and Credit. Preliminary analyses were performed separately for the combined data from the two

Table 3
Correlations Among Final Verdict and Positions on Five Points of Dispute: Experiment 3

Point of dispute	1	2	3	4	5	6
			Pretest			
1. Truth	—	.11	−.04	−.06	.10	−.02
2. Cause		—	−.11	.05	.10	.05
3. Regulation			—	.06	−.14	.09
4. Speech				—	.51**	.27*
5. Analogy					—	.11
6. Verdict						—
			Posttest			
1. Truth	—	.80**	.49**	.28*	.37**	.56**
2. Cause		—	.57**	.38**	.45**	.67**
3. Regulation			—	.27*	.45**	.73**
4. Speech				—	.52**	.43**
5. Analogy					—	.53**
6. Verdict						—

*$p < .05$. **$p < .01$.

experimental conditions and for the baseline condition that did not receive the Quest case. As the two sets of analyses yielded very similar results, we report a single analysis based on all 80 participants in the study, collapsing over the three conditions. We performed an ANOVA to determine how assessments of the two points of dispute changed across tests as a function of whether the verdict favored the side citing the newspaper versus the telephone analogy. (Recall that given the counterbalancing used in the study, this division is independent of whether the verdict was for Infoscience or for the employees.) The dependent measure was Q-score, defined as favorability to the position of the side that cohered with Quest (i.e., the side that agreed with the newspaper analogy).

Figure 8 shows the shifts in mean Q-scores from the pretest to the posttest as a function of the verdict in the Infoscience case (i.e., whether the winning side cited the newspaper or the telephone analogy). The interaction apparent in Figure 8 was highly significant, $F(1, 78) = 36.4$, $MSE = 2.36$, $p < .001$, and did not differ significantly between the two points of dispute, $F(1, 78) = 2.72$, $MSE = 1.80$, $p > .10$. Tests of simple main effects revealed that the mean Q-score was somewhat higher for those who eventually chose the side citing the newspaper analogy as opposed to the telephone analogy (.66 versus .06, $p < .05$); however, this difference was much larger on the posttest (1.19 versus -1.52, $p < .001$). The correlation between positions on Analogy and Credit was not significant on the pretest, $r(76) = .10$, $p > .30$, but was strongly positive on the posttest, $r(76) = .42$, $p < .001$. Within the Analogy factor, the pretest correlation between ratings of the two analogies was nonsignificant, $r(76) = -.15$, whereas the comparable posttest correlation was $r(76) = -.48$, $p < .001$. Thus the Infoscience case, like the Quest case, displayed the pattern

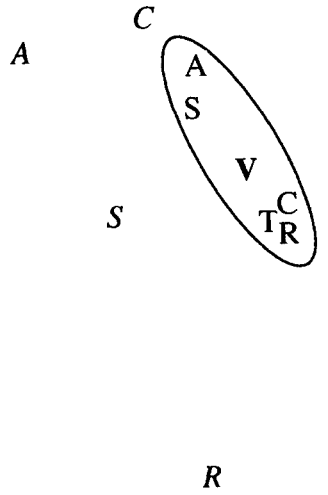


Figure 7. Two-dimensional nonmetric scaling solution for correlation matrix among positions on the disputed points in Experiment 3, displaying the shift from unrelated initial positions on the pretest (italicized) to a coherent position associated with the verdict on the posttest (Roman type). V = verdict; A = analogy; C = cause; M = motive; R = regulation; S = speech; T = truth.

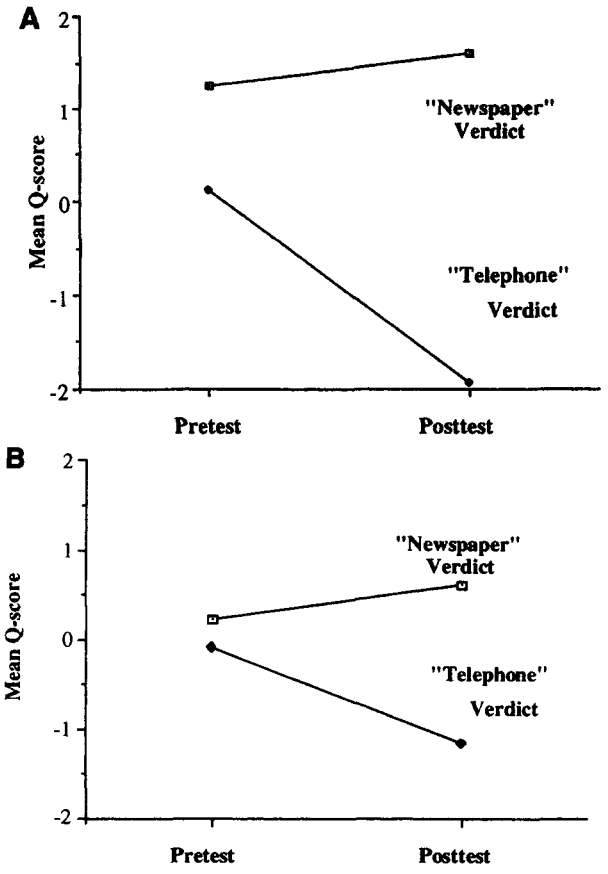


Figure 8. Shifts in Q-scores (favorability to position cohering with that of Quest) across tests for the Infoscience case, as a function of eventual verdict for the side citing the newspaper (B) versus telephone analogy (A) in Experiment 3. Lower scores indicate less favorability to the position cohering with that of Quest, whereas higher scores indicate more favorability to it.

of internal coherence predicted by coherence-based models of decision making.

Transfer from Quest case to Infoscience case. We performed a further series of analyses to determine whether participants' judgments in the Infoscience case were influenced by a tendency to maintain coherence with the Quest case. The percentage of participants who decided in favor of the Infoscience side that cited the newspaper (i.e., pro-Quest) analogy tended to be higher in the Bad Smith than the Good Smith condition (66 versus 50) and higher among those who decided in favor of Quest than among those who decided in favor of Smith (72 versus 44), with the percentage for participants in the baseline condition (who did not receive the Quest case) falling in between for both comparisons (63). However, both of the above comparisons fell short of statistical significance, $\chi^2(1, N = 64) = 1.60$ and 2.87, respectively, $p > .10$.

Nonetheless, we identified a number of statistical links between the two cases. The critical bridge—the link between the assessment of Analogy in the context of the Quest case and its assessment in the context of the Infoscience case—

was highly significant, $r(62) = .45$, $p < .001$. In addition, Analogy within the Quest case cohered with the Infoscience verdict, $t(62) = 3.81$, $p < .001$, and the Quest verdict cohered with Analogy within the Infoscience case, $t(62) = 2.21$, $p < .05$. Perhaps most remarkably, the experimental manipulation of Smith's character in the Quest case influenced the Q-score for Credit in the Infoscience case. Participants in the Bad Smith condition had a higher Q-score on Credit than did those in the Good Smith condition (.20 versus −.48), $t(62) = 2.01$, $p < .05$. Thus, the longest inference path along which coherence could spread (see Figure 9) proved to be statistically reliable.

Structural equation modeling. In order to assess the adequacy of the overall coherence model as a description of the causal connections among the manipulation of Smith's history, the assessments of arguments, and the verdicts, we formulated the model as a structural equation model. We used the EQS program (Bentler, 1995) to estimate the model's parameters, test their significance, and assess the global fit and parsimony of the model. Figure 9 provides a graphical display of the initial and final coherence models that were tested. Directed arrows connect predictor variables to their outcomes. The arrows follow the major predicted flow of influence from the experimentally manipulated variable, Smith's history; we did not attempt to estimate separate bidirectional influences because the number of additional parameters required would have been excessive relative to the number of data points being fitted. As all variables were measured, no latent variables were included in the model. The direct predictive links are from Smith's history (a binary independent variable) to Motive; to the Quest verdict; to each of the other Quest arguments, including Analogy; from Analogy in Quest to Analogy in Infoscience; to the Infoscience verdict; to Credit (a total of six inference steps). Each argument variable was represented by its Q-score (a continuous measure). To maximize normality of the variables, the two verdicts were also represented as continuous measures, formed by weighting the verdict by its confidence (e.g., a verdict for Quest with a confidence rating of 5 yielded the maximal Quest verdict of 5; a verdict for Smith with a confidence rating of 5 yielded the minimal Quest verdict of −5). We used structural equation modeling to assess overall model fits, to estimate a parameter value for each direct predictive link, and to estimate standard errors of predictions.

The independence model that tested the hypothesis that the variables were uncorrelated with one another was easily rejected, $\chi^2(55, N = 64) = 414, p < .001$. Then we tested the initial coherence model. Because the variables of the model deviated somewhat from normality (Mardia's coefficient normalized estimate = 2.33), maximum likelihood estimation with the Satorra-Bentler scaled chi-square was used in all analyses. A chi-square difference test indicated that the initial coherence model provided a significant improvement over the independence model ($p < .001$). The initial coherence model provided a fairly good fit to the observed pattern of assessments and verdicts, although its deviations were significant, Satorra-Bentler $\chi^2(45, N = 64) = 83.7, p < .001$. The robust Comparative Fit Index (CFI) was .90 (where CFI can range from 0 to 1 and values over .90 are

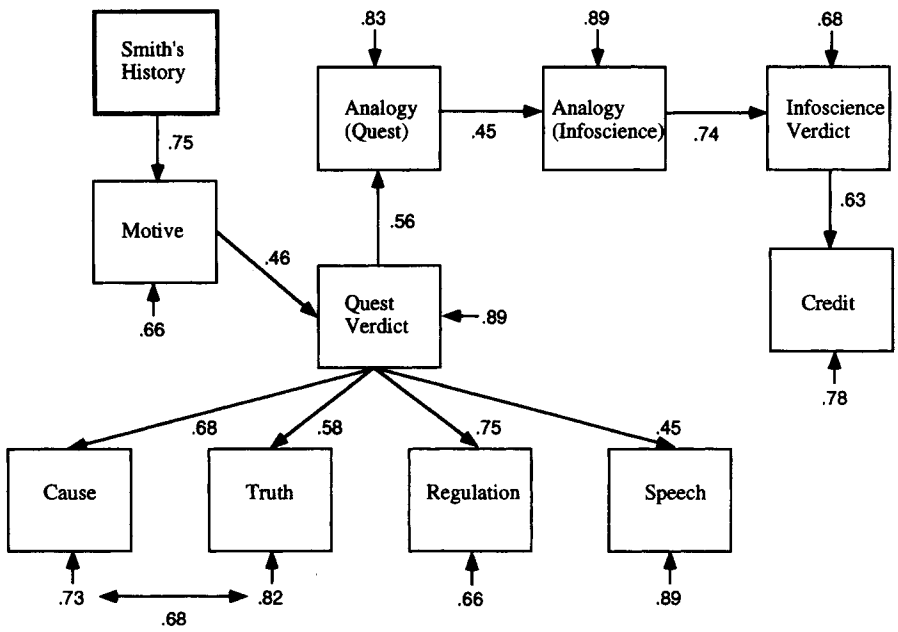


Figure 9. Graphical summary of the final model of coherence shifts among the experimental manipulation of Smith's character, verdicts, and points of dispute in the Quest and Infoscience cases in Experiment 3, with parameter estimates and error terms derived from EQS. The final model differed from the initial model solely by the addition of a parameter reflecting the correlated errors associated with the Cause and Truth factors.

typically considered indicative of good-fitting models; Ullman, 1996). The Lisrel goodness-of-fit index (GFI; analogous to variance accounted for) was .81. The Akaike Information Criterion (AIC), an index that takes into account the parsimony of the model (i.e., inverse of number of parameters) as well as its fit to the data, was 6.25 (where low scores indicate a parsimonious and accurate fit).

The Lagrange Multiplier Test suggested the initial model could be improved by estimating an additional parameter to reflect the correlation between the residual errors associated with Cause and Truth. This augmentation in the model was quite sensible, as assessments of these two factors reflect closely linked aspects of Smith's message. It could be, for example, that people who believed the message was true were also likely to believe that poor management, rather than the message itself, was the cause of Quest's downfall. This additional link between Cause and Truth is included in Figure 9, which also provides the parameter estimates and error terms derived from this final coherence model. The final model was a significant improvement over the initial model, Satorra-Bentler $\chi^2(1, N = 64) = 31.7, p < .001$, and provided an excellent fit to the data; indeed, the final model could not be statistically rejected, Satorra-Bentler $\chi^2(44, N = 64) = 51.4, p = .21$. The robust CFI was .98, the GFI was .88, and the AIC was −30.3.

The ten parameter estimates for weights on direct links were identical in the initial and final models, $r(8) = 1.0, p < .01$, increasing confidence in their stability. The parameter values on the direct links in Figure 9, taken from the final model, are standardized values representing the predicted shift in the outcome variable in standard deviation units as a function of a one standard deviation shift in the predictor variable (controlling for the influence of any other predictors). Estimates of residual error for each outcome variable are also shown in Figure 9. All direct links between predictor and outcome variables (as well as the undirected link between the residual errors for Cause and Truth) were significant, $p < .01$. EQS also assesses the significance of indirect effects of predictor variables on outcome variables separated by one or more mediating variables. The strength of an indirect link is a function of the product of the strengths of the direct links along the chain (e.g., the estimated indirect effect of Smith's history on the Quest verdict, which is mediated by Motive, is $.75 \times .46 = .35$). The main question of interest concerned how extensive were the coherence effects triggered by Smith's history (the independent variable in Experiment 3). The results were clear: z tests revealed that *all* the indirect effects of Smith's history (indeed, all indirect effects in the entire final model) were significant, including the most remote, 6-step indirect influence of Smith's history on Credit in the Infoscience case. The estimated strength of the indirect influence of Smith's history across the full 6-step chain (given by the product of the six links on the chain) was small, .04, but reliable, $p < .05$. Structural equation modeling thus provided further support for the claim that our experimental manipulation of Smith's character generated coherence shifts that extended throughout both the Quest and the Infoscience cases.

We also used structural equation modeling to assess a class of potential alternative accounts of some of the coherence effects observed in Experiment 3. Our coherence model assumed that the sole direct effect of the manipulation of Smith's character was its impact on the assessment of Motive, which then in turn had a direct effect on the Quest verdict, which in turn had direct effects on the assessments of the other Quest arguments. It might be argued that the manipulation of Smith's character instead had more widespread direct effects by virtue of its impact on the defendant's credibility. That is, if Smith had a history of fraud, his defense may have been generally viewed as less credible than if he had a history of scrupulous honesty. This credibility account would thus explain much of the observed pattern of associations between variables as being attributable not to spreading coherence but, rather, to the direct impact of Smith's history on the assessment of multiple Quest arguments.

The credibility account was specified and tested using EQS. We used three variants of the credibility account. Credibility Model 1 assumed that Smith's history directly predicted assessments of all the arguments of a factual nature related to Smith's actions. Relative to the initial coherence model (see Figure 9), Credibility Model 1 added direct links from Smith's history to Cause, Truth, and Regulation (in addition to Motive); each of these four arguments then predicted the Quest verdict. Credibility Model 2 assumed that Smith's credibility directly affected not only assessments of the factual arguments, but also assessments of the more general legal issues, Speech and Analogy. Hence, direct links were added from Smith's history to these two variables, which then were also used to predict the Quest verdict. In Credibility Model 2, therefore, Analogy (like all the other arguments) was a predictor of the Quest verdict, rather than vice versa (as in the coherence models). It might be argued that if Credibility Model 2 did not fit well, it could be due to the lack of a predictive link from the Quest verdict to Analogy (which in turn would then predict the assessment of Analogy in the Infoscience case). To assess this possibility, we constructed Credibility Model 3 to be identical to Model 2 except that the direction of the link between the Quest verdict and Analogy was reversed, so that the former predicted the latter. Note that (unlike the final coherence model) none of the credibility models included an explicit link between the errors associated with Cause and Truth. However, all of the credibility models included direct links from Smith's history to both Cause and Truth—links that could potentially account for the high correlation observed between these two factors.

In fact, however, none of the three credibility models was able to provide a satisfactory fit to the data. As these models were not nested with respect to the coherence models, the two classes of models could not be directly compared by difference chi-squares. However, all indirect comparisons favored the final coherence model (and also the initial coherence model) over any credibility model. The best fitting credibility model was Model 1, in which Smith's history had direct links to Motive, Cause, Truth, and Regulation. Whereas the final coherence model could not be statistically rejected as an account of the data, Credibility

Model 1 (as well as Models 2 and 3) could be, Satorra-Bentler $\chi^2(42, N = 64) = 101.5, p < .001$. Relative to the coherence models, Credibility Model 1 also had a lower robust CFI, .85, a lower GFI, .78, and a higher AIC, 32.3, indicating a poorer and less parsimonious fit (as all the credibility models required more free parameters than did the coherence models).

In sum, structural equation modeling provided strong confirmatory evidence favoring the coherence models in terms of both their absolute fit and their fit relative to an alternative class of models. In particular, the influence of the experimental manipulation of Smith's character was better captured by viewing its direct effect as restricted to Motive (with indirect effects spreading throughout the network as indicated in Figure 9), rather than by viewing the manipulation as having direct effects on assessments of other arguments in the Quest case.

Recall of Initial Positions

As in Experiment 2, we assessed participants' memory for their pretest ratings by computing Q-scores based on the recall test. The analyses to be reported involve the experimental conditions only. One participant failed to complete the memory test and therefore was not included in these analyses. One set of analyses involved mean Q-scores for the Quest-related factors other than Motive (i.e., Truth, Cause, Regulation, Speech, and Analogy). The mean recalled Q-score was .16 for the Quest supporters, higher than their actual mean pretest Q-score of −.10; the mean recalled score for the Smith supporters, −.81, was lower than their actual mean score of −.34. The difference between the recalled Q-scores for the two subgroups was significantly larger than the actual pretest difference, $F(1, 61) = 7.27$, $MSE = 0.58$, $p < .01$; the two subgroups did not differ reliably on the pretest, $p > .40$, but did differ in their recalled positions, $p < .001$. Across individual participants, the recalled Q-score was correlated with the posttest Q-score, $r(61) = .57, p < .01$, the actual pretest Q-score, $r(61) = .52$, $p < .01$, and the verdict on the Quest case, $r(61) = .44, p < .01$. A stepwise multiple regression analysis revealed that all three predictor factors contributed significantly to prediction of the recalled Q-score, with the posttest Q-score being the strongest single predictor.

We performed a parallel set of analyses on memory for pretest position on Credit, the one factor unique to the Infoscience case. Mean recalled Q-score for Credit did not differ as a function of the verdict on the Infoscience case (i.e., whether the favored side cited the newspaper or the telephone analogy). A correlational analysis revealed that recalled Q-score for Credit was reliably predicted by the posttest Q-score for Credit, $r(61) = .47, p < .01$, and by the pretest Q-score, $r(61) = .29$, $p < .01$, but not by the Infoscience verdict, $r(61) = .11$, $p > .20$. A regression analysis confirmed that both the pretest and the posttest Q-scores contributed to prediction of the recalled Q-score, with the posttest Q-score being the stronger predictor.

In general, then, these analyses indicated that recall of initial positions was biased by the intervening decision: The recalled position was, in essence, a blend of the actual initial position and the position on the posttest. With a relatively short delay between the pretest and the recall test in Experiment 3, the best single predictor of the recalled positions was, in fact, the person's position on the posttest, rather than his or her actual pretest ratings.

Simulation of Coherence-Based Reasoning With Co3

Although we interpreted our findings in terms of coherence-based decision making, it would be desirable to show by simulation that such coherence shifts could in fact be generated by a computational model operating by constraint satisfaction. As Experiment 3 provided the most complete data on decision making and transfer by spreading coherence, we performed a simulation of it using a variant of ECHO (Thagard, 1989, 1992) called Co3 (Coherence Model of Cognitive Consistency; Spellman et al., 1993). Co3 is based on the ECHO program but does not enforce all of the constraints of Thagard's (1989) theory of explanatory coherence (e.g., simplicity is not considered, and excitatory connections can be based on looser conceptual relations). Co3 allows networks to be formed sequentially, saving prior states of the network as new units and links are added. It also allows attentional focus to be modeled by the probability that any unit will be active (i.e., will produce an output and have its activation updated). The Co3 simulation is presented as a sufficiency proof that a constraint satisfaction algorithm can generate the qualitative pattern of coherence shifts observed in our experiments. Other similar models (e.g., Kunda & Thagard, 1996; Shultz & Lepper, 1996) could also be adapted to simulate our findings. We did not attempt to generate quantitative fits to our data; rather, we simulated the coherence shifts of a single hypothetical reasoner using Co3 as a concrete instantiation of a constraint satisfaction model. (See Ranney & Thagard, 1988, and Schank & Ranney, 1991, for similar applications of ECHO as a model of the revision of beliefs as new evidence accumulates over time.) Like other models of this nature, Co3 only simulates the evaluation of a network of interconnected arguments, rather than the processes by which the network itself is actually constructed (see Kunda & Thagard, 1996). (For models and evidence concerning the construction of explanations, see Pennington & Hastie, 1986, 1988, 1994; Read & Miller, 1993.) In the present simulations, we simply assumed the general topology of the network that we believed a participant would typically construct. The performance of the model was a joint function of the assumptions we made about the connectivity of the constructed network and of the constraint satisfaction process that then operated on it. The nature of the process of network construction is clearly an area in which further modeling work is required.

Figure 10 depicts the state of the network at three stages corresponding to the pretest (Panel A), posttest on the Quest case (Panel B), and transfer to the Infoscience case (Panel C). Each unit represents an argument or a possible verdict, with a pair of units representing opposing sides on each point of dispute in the two cases (e.g., A–news represents the proposition "the Internet is analogous to a newspaper,"

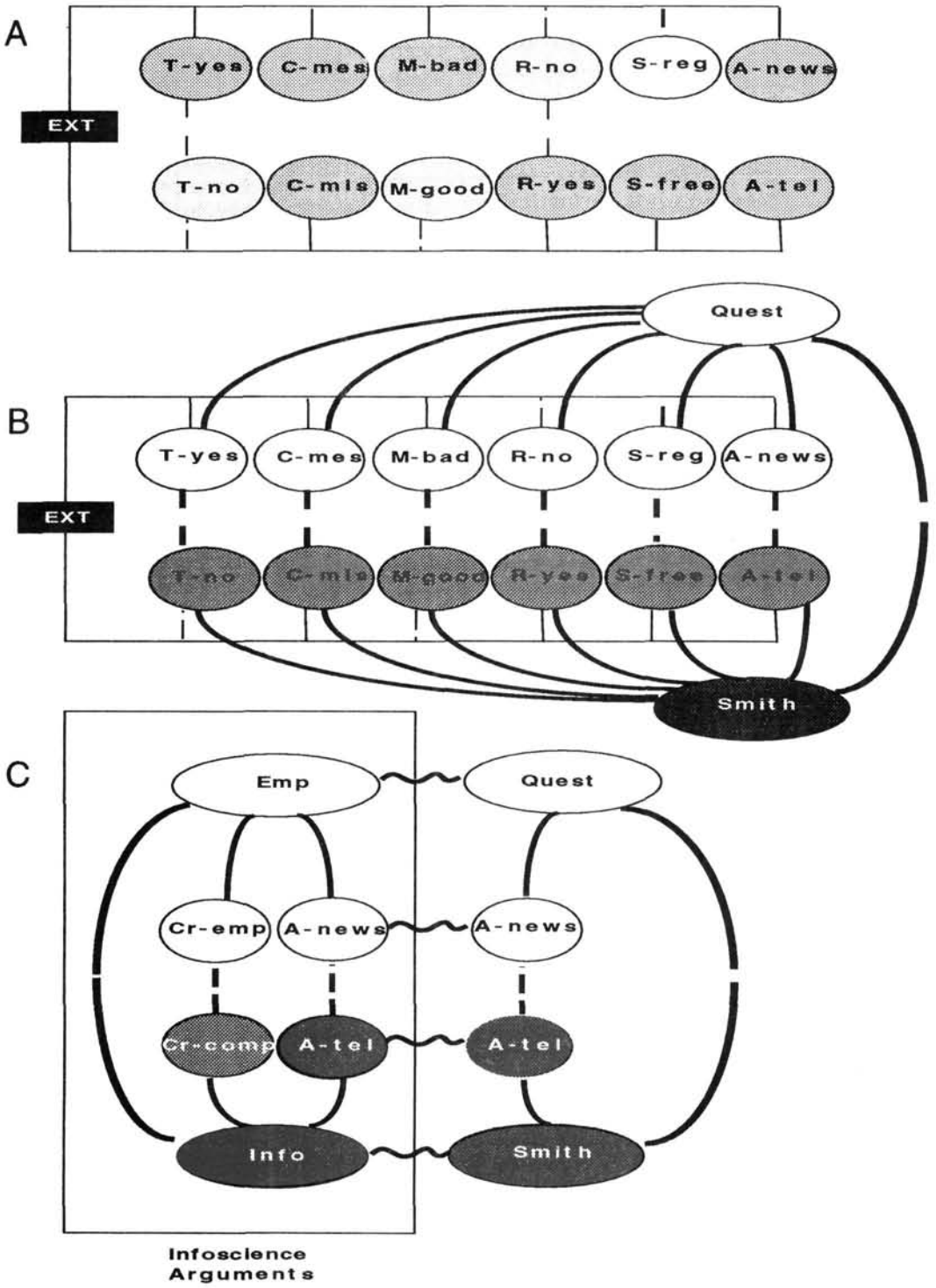
whereas A–tel represents "the Internet is analogous to a telephone system"). Excitatory links (solid lines) connect propositions that support one another (e.g., in Panel B, A–news supports a verdict for Quest), and inhibitory links (dashed lines) connect propositions that conflict. The units representing propositions are linked to a special external (EXT) unit, which is clamped on and continually passes activation (either excitatory or inhibitory) to each unit connected to it. The EXT unit represents prior knowledge that influences judgments about individual arguments.

Panel A represents the initial network for the hypothetical reasoner, representing the relatively unstructured knowledge state prior to encountering the Quest case. The individual units are generally not linked to one another, except for the salient incompatibilities between whether the message is true or not (T–yes and T–no) and whether Smith complied with the company regulation or not (R–yes and R–no). (The inhibitory weights are both −.01.) The reasoner is assumed to have weak entering opinions based on prior knowledge (i.e., links from the EXT unit) that influence the pretest judgments about the various arguments. The links are either .01 or −.01 (except a −.02 link to S–reg, indicating slightly stronger disagreement with Internet regulation). It is important to note that all links represent soft constraints, that is, preferences rather than requirements. This network was allowed to settle for 50 cycles of updating, using McClelland and Rumelhart's (1981) algorithm with decay parameter set at .05. (See Spellman et al., 1993, or Thagard, 1989, 1992, for a more detailed description of the constraint satisfaction process.) The shading of the units represents the resulting activations, where white indicates a distinctly negative activation (<−.25) and darker grays indicate increasingly positive activation, with a maximal activation of 1. As Panel A reveals, the resulting activations are weak (ranging from −.28 to .18) and generally uncorrelated with the eventual linkage to the possible Quest verdicts. The reasoner has a positive assessment of four arguments that will eventually favor Quest (top row) and four that will favor Smith (bottom row). (The slightly more negative assessment of S–reg [−.28] than any other argument provides the tiny seed from which the reasoner's verdict will eventually grow in Panel B.) The simulation captures a typical qualitative pattern of pretest assessments—relatively neutral evaluations, with little correlation with each other or the eventual verdict.

Panel B depicts the revised network that we assume is created as the reasoner processes the facts and arguments introduced in the Quest case. Each argument now has a strong (.10) excitatory connection to the verdict it supports and a strong (−.10) inhibitory connection to its direct rival. The two verdicts inhibit each other very strongly (−.20), thus enforcing a winner-take-all outcome. The entering biases (connections to EXT) remain unchanged from Panel A, and the final activations in Panel A become the entering activations used in simulating posttest assessments. The shading in Panel B indicates the activations of the units after an additional 50 cycles of updating. The hypothetical reasoner has firmly decided in favor of Smith (activation of .92) and has accepted all the arguments in favor of Smith's position while strongly rejecting all the arguments in favor of Quest's position. As noted above, the seed for the decision is the reasoner's initial disagreement with S–reg, the argument that the Internet should be regulated. Although this hypothetical reasoner is more likely to decide for Smith than Quest, Co3 allows stochastic updating and thus models the reasoner's decision as probabilistic, rather than as strictly determined by prior beliefs. (That is, other runs with the same network will yield the opposite verdict, with the argument assessments also reversed to maintain coherence.) The simulation captures the basic coherence shift—the internal structure imposed by the case allows constraint satisfaction to change the reasoner's assessments from a weak, uncorrelated muddle (Panel A) to a strong, internally coherent system (Panel B).

Panel C depicts the further development of the network that is triggered when the Infoscience case is presented. We assumed the reasoner notices the continued relevance of the two rival analogies, A–news and A–tel. We adopted ECHO's use of analogy as a constraint (Thagard, 1989), generating excitatory connections (wavy lines, with weights of .10) across the two cases between analogous arguments and the verdicts they support. Thus, A–tel in the Quest case is analogous to A–tel in Infoscience, and the Smith verdict (supported by A–tel) is analogous to a verdict for Infoscience (modeling the situation in which Infoscience argues that an Internet company is similar to a telephone company). For simplicity, we assumed our hypothetical reasoner had no prior opinions about the Infoscience arguments; hence the eventual verdict was determined only by the analogies. (We also ran the simulation with weak links to EXT supporting the arguments favoring a verdict for the employees; the accepted analogy remained the decisive factor.) In deciding the Infoscience case, the reasoner presumably attends directly only to the arguments made in the Infoscience case itself, plus the relevant analogies from the Quest case. In the simulation, therefore, all the other Quest units (not shown in Panel C) were shut down (i.e., did not participate further in constraint satisfaction). After 50 further cycles of updating, the activations in Panel C show that coherence has spread from the Quest case to the Infoscience case. Because the reasoner has come to accept the telephone analogy in the course of deciding the Quest case, analogical transfer provides support for the telephone analogy in the Infoscience case, which in turn supports a verdict for Infoscience and (slightly less strongly) assignment of credit for profits to the new computers, rather than the efforts of the employees. Note that the telephone analogy controls the eventual Infoscience verdict (Panel C) even though the reasoner was initially equally positive toward both the telephone and the newspaper analogies (Panel A).

In sum, the Co3 simulations provided a demonstration that a constraint satisfaction algorithm can produce the qualitative pattern of coherence shifts obtained in Experiment 3 and the earlier experiments. Experiment 3 provided direct experimental support for the assumption that analogy provides a source of constraint in evaluating competing options (Thagard, 1989).


A
EXT
T-yes
C-mes
M-bad
R-no
S-reg
A-news
T-no
C-mis
M-good
R-yes
S-free
A-tel
B
Quest
EXT
T-yes
C-mes
M-bad
R-no
S-reg
A-news
T-no
C-mis
M-good
R-yes
S-free
A-tel
Smith
C
Emp
Quest
Cr-emp
A-news
A-news
Cr-comp
A-tel
A-tel
Info
Smith
Infoscience
Arguments

General Discussion

The Emergence of Coherence During Decision Making

The present findings provide strong support for coherence-based models, which treat the evaluation of rival arguments and decisions as a process of soft constraint satisfaction based on a connectionist-style network (e.g., Holyoak & Thagard, 1989; Kunda & Thagard, 1996; Shultz & Lepper, 1996; Spellman et al., 1993; Thagard, 1989, 1992; Thagard & Millgram, 1995). In all three experiments, the assessments of multiple arguments underwent a profound shift triggered by the presentation of a legal case in which the previously unrelated arguments were organized into the opposing positions of the plaintiff and defendant. Prior to presentation of the case, participants' assessments of the various points were generally uncorrelated, both with one another and with the verdict that would eventually be reached. The only consistent pretest predictor of the eventual verdict (significant in Experiments 1 and 3) was the assessment of a statement asserting that "communications over the Internet ought to be regulated by law," which elicited greater support from participants who would eventually decide that the plaintiff company had been libeled than from those who would decide in favor of the defendant. After participants had studied and evaluated the case, their assessments of all the disputed points had shifted so as to cohere both with one another and with the verdict. The basic coherence shift was obtained both within a single session (Experiments 1 and 3) and when a 1-week delay intervened between the pretest and the process of deciding the case (Experiment 2). Despite the apparent ambiguity and conflict inherent in the case, participants reported high levels of confidence in their decisions. In Experiment 1, for example, 75% of participants reported maximal or next-to-maximal confidence that their decision was the best possible, and only 5% reported low confidence.

The comparison of the 2-phase and 3-phase conditions of Experiment 1 provided evidence that the shifts in participants' assessments of arguments was not due simply to a post hoc reduction in dissonance in the aftermath of a decision that had been reached for some other reason (Festinger, 1964). Participants in the 3-phase condition made preliminary assessments of the arguments after having been advised to withhold final judgment pending receipt of additional relevant information (a decision to be handed down in a related case). The fact that about 20% of these participants would later change their verdict (even though no new information was actually presented) provides confirmatory evidence that the 3-phase participants did not always commit themselves to a verdict prior to the interim test. The results clearly showed that essentially the entire coherence shift occurred between the pretest and the interim test, and that coherence was then maintained between the interim and final tests (except for those participants who switched sides from their preliminary leaning to their final verdict). These findings support the hypothesis that the emergence of a coherent position guides the process of decision making and thus has a causal impact on the eventual verdict.

Experiment 3 demonstrated that broad and far-reaching coherence shifts could be experimentally manipulated. (See Spellman & Holyoak, 1992, 1996, for similar manipulations of coherence in analogical mappings.) By providing differential information about the character of the defendant, we were able to influence participants' verdicts. Moreover, the manipulation triggered shifts in assessments of points of dispute that were only indirectly linked to the character of the defendant (e.g., if the defendant was characterized as having a history of unscrupulous behavior, participants became more supportive of regulation of Internet transmissions and less supportive of electronic free speech). Structural equation modeling showed that the coherence model provided a more accurate and more parsimonious fit to the data than did alternative models in which the credibility of the defendant was assumed to have a direct impact on

Figure 10 (opposite). Networks showing Coherence Model of Cognitive Consistency (Co3) simulation of a hypothetical reasoner who decides Quest case and then Infoscience case. Units represent arguments, and shading represents activation of unit (white: $<-.25$; black: 1). Solid lines are excitatory links, and dotted lines are inhibitory links. Thickness of line represents absolute magnitude of weight on link. Wavy lines are excitatory connections based on analogy. Short lines connected to each unit are links to external (EXT) unit, which represents prior beliefs that influence assessments of arguments. T–yes = "message is true"; T–no = "message is false"; C–mes = "message caused Quest's bankruptcy"; C–mis = "mismanagement caused Quest's bankruptcy"; M–bad = "Smith's motive was vindictiveness"; M–good = "Smith's motive was altruistic"; R–no = "Smith violated company regulation"; R–yes = "Smith complied with company regulation"; S–reg = "speech on Internet should be regulated"; S–free = "speech on Internet should be free"; A–news = "the Internet is analogous to a newspaper"; A–tel = "the Internet is analogous to a telephone system"; Quest = "verdict is for Quest"; Smith = "verdict is for Smith"; Cr–emp = "credit for profits is due to employees"; Cr–comp = "credit for profits is due to new computers"; Emp = "verdict is for employees"; Info = "verdict is for Infoscience." Figure 10A shows initial network structure and activations after 50 cycles (simulating pretest ratings). Figure 10B shows network structure after reading Quest case and activations after an additional 50 cycles (simulating posttest ratings). Figure 10C shows active portion of Quest network and Infoscience network (in box), with activations after a further 50 cycles (simulating transfer from the Quest decision to the Infoscience case).

multiple arguments in the Quest case. We used the Co3 model to provide a demonstration that the spread of coherence can be modeled by a process of parallel constraint satisfaction.

Similar indirect influences of changed attitudes were observed by Spellman et al. (1993) in their observational study of opinion change during the Persian Gulf War and were also simulated by the Co3 model. For example, people who came to view Saddam Hussein (the leader of Iraq) more negatively also tended to display diminished support for pacifism, even though these two attitudes lacked any clear logical connection. Although both the present study and that of Spellman et al. (1993) demonstrated coherence shifts, there are some major differences between the basis for the shifts observed in the two studies. First, Spellman et al.'s study focused on attitudinal links, whereas the present study focused more directly on inferential links between arguments and a decision. Second, Spellman et al.'s study was entirely observational, whereas Experiment 3 of the present study introduced an experimental manipulation that altered coherence patterns. Third, the shifts in attitudes observed by Spellman et al. appeared to result from perturbations of particular beliefs connected within a stable network of linked attitudes. Both before and after the attitude shifts observed by Spellman et al., participants' attitudes were correlated with one another. In contrast, the present study first measured participants' evaluations of disputed points before a densely connected constraint network was established, at which point the evaluations were generally uncorrelated.

The present findings, and particularly the transfer results obtained in Experiment 3, lend support to McGuire's (1960, 1990) conception of cognitive inertia, according to which changes in one thought reverberate through a loosely linked chain to influence related thoughts. More generally, our findings support McGuire's theory of the dynamic operations of thought systems (McGuire, 1990). This theory focuses on thought systems that revolve around a core event, including assessments of that event's desirability and likelihood of occurrence, as well as its antecedent causes and its consequences. McGuire (1981, 1990) showed that changes affected by persuasion in any one of these aspects of the representation of a core event can generate changes in the other components.

The present findings are also compatible with Pennington and Hastie's (1986, 1988, 1994) story model of juror decision making, which describes how the organization of trial evidence into a coherent story narrative influences jurors' decisions. The story model describes the emergence of newly created belief networks and the processes used to reason toward a decision (a criminal verdict). The model is primarily designed to answer the causal question "What happened?" and the cases studied by Pennington and Hastie focused on inferences of physical and intentional causality. Our results are also in accord with Thagard and Millgram's (1995) goal-based theory of deliberative coherence, according to which people make decisions by assessing the coherence of a best plan given the person's goals and possible actions. In integrating these factors, the person's goals may undergo reevaluation and even rejection. The decision task in the present study involved a broader range of inferences than those discussed by either Pennington and Hastie or Thagard and Millgram, including assessment of causality, evaluation of analogies, assessment of compliance with a proscriptive regulation, and application of public policy considerations.

Analogy as a Conduit of Spreading Coherence

One of the most important argument types involved in coherence-based decision making is analogy. It has been argued that analogical mapping (Holyoak & Thagard, 1989) and retrieval (Thagard, Holyoak, Nelson, & Gochfeld, 1990) are themselves based on constraint satisfaction; moreover, analogy has been hypothesized to play a role in constraining the assessment of competing explanations (Thagard, 1989, 1992). In the present study, one point of dispute involved a pair of competing analogies: For legal purposes, should messages over the Internet be likened to articles in a newspaper (which are subject to libel law) or to telephone calls (which are not)? Assessments of these analogies shifted in accord with the other disputed points and the verdict, as each analogy simultaneously supported and was supported by the other points constituting a coherent position. Thus, on the pretest, the evaluation of each analogy tended to be mildly positive and the assessments were generally uncorrelated; however, after the case had been presented, each assessment shifted up or down to cohere with the verdict, and hence the assessments became negatively correlated with one another.

Analogy is a powerful source of transfer because a rich analogy can potentially support an indefinitely large number of inferences. In Experiment 3, the analogy that emerged as the winner in the initial case served as a bridge, yielding transfer of coherence across to a second case. For example, a participant who decided for Quest was likely to end up favoring the newspaper analogy, which supported a finding that Smith libeled Quest. Once the newspaper analogy came to be favored as the better match to the Internet, the participant was also likely to use the same analogy to decide how large a bonus the employees of an Internet company should receive (namely, a bonus similar to that received by employees of a newspaper rather than that received by employees of a telephone company). As coherence continued to spread, the decision in the second case also molded assessments of a yet more remote point of dispute—whether the company's profits were attributable to the efforts of the employees or to a newly installed computer system—a point of dispute with no logical connection to either analogy.

The form of analogical transfer identified in Experiment 3 differs from those emphasized in previous studies of analogy. Earlier work has shown that alternative analogies compete to be retrieved from memory (Wharton et al., 1994; Wharton, Holyoak, & Lange, 1996), and that within a single retrieved analogy there is competition among alternative correspondences between its elements (Spellman & Holyoak, 1992, 1996). However, the present form of competition operates between two analogies after both have been re-

trieved from long-term memory. There is reason to believe that the present form of interanalogy competition is especially important in argumentation. In debates about the merits of American intervention in the Persian Gulf crisis of 1990–1991, the two sides did not argue about details of a mapping (e.g., whether Saddam Hussein corresponded to Hitler or to Mussolini) but, rather, about what overall analogy was most appropriate (e.g., whether the Gulf situation was more like that in Europe just prior to World War II or like that in Vietnam just prior to American intervention there). Such interanalogy competition has been observed in case studies of the use of analogy in political debates, such as American decision making during the Vietnam War (Khong, 1992) and arguments in Canada concerning Quebec separatism (Blanchette & Dunbar, 1997).

Although the selection of one analogy among a small set of competitors differs from analogical mapping within a single analogy, the two processes may be theoretically linked. In fact, Holyoak and Thagard (1989) showed that the ACME model of mapping predicts that when people are faced with a choice between two analogies that are approximately equal with respect to structural and semantic constraints, they will prefer the analogy that supports pragmatically desirable inferences. ACME would map the Nicaraguan "contras" of the 1980s onto "freedom fighters" if the model had a slight bias to prefer that the United States should support the contras but onto "terrorists" if the model had the opposite preference. To model the impact of goals on interanalogy competition, ACME mapped the target analog (contras) onto a source analog defined as the union of the representations for freedom fighters and for terrorists. Thus, both intra-analogy and interanalogy competition may be governed by the same general constraints on analogical mapping. In turn, interanalogy competition provides one source of constraint on coherence-based decision making.

Impact of Coherence on Memory for Initial Positions

An additional type of coherence effect observed in the present study (weakly in Experiment 2 and robustly in Experiment 3) involved participants' recall of their initial assessments of the various arguments. Once participants had reached their final decisions, they had difficulty in recalling the initial inferences that they had reported on the pretest. The inferences recalled were somewhere in between the initial ones and those they reported on the posttest, with the latter being the single best predictor. The result is an apparent tendency to believe that whatever one now believes, one has always (to some degree) believed. To model such memory biases, constraint satisfaction models, such as Co3, would need to be augmented with the capacity to store more permanent, but nonetheless malleable, representations of prior states of the belief network.

The memory shift observed in the present study may be related to a variety of similar phenomena in the literature. After undergoing a change of attitude, people tend to have weak recall of their former attitudes (Bem & McConnell, 1970; Goethals & Reckman, 1973; Ross & Shulman, 1973; Wixon & Laird, 1976). In a similar manner, the hindsight bias is a general tendency to view actual outcomes as having been more foreseeable than was actually the case (Fischhoff, 1982). Other memory studies have shown that interventions after the initial experience can bias subsequent recall (e.g., Higgins & Liberman, 1994; Loftus & Palmer, 1974). The present study raises the possibility that at least some of these apparently similar memory biases may reflect the impact of cognitive mechanisms that tend to maintain coherence of beliefs—both coherence among the beliefs a person holds at a single time and (partly illusory) coherence among the beliefs a person holds across different times.

Further Implications

The most general conclusion of the present study is that it is possible to create dependencies among beliefs, attitudes, opinions, and decisions by introducing a set of links connecting individual variables to a common outcome. To the extent a particular outcome comes to be favored, synergistic feedback loops will tend to generate a coherent position across all the interconnected variables. The common assumption that inferences are unidirectional, flowing from premises accepted as true to conclusions, is thereby weakened. Bidirectional links not only allow premises to support belief in their conclusions, but also allow conclusions to support belief in premises.

Although the present results support the importance of bidirectional links in decision making, we do not mean to deny that some inferences are strongly directional. For example, given that it is raining, we can securely conclude that the lawn will be wet; however, given that the lawn is wet (at least in southern California), it would be unwise to conclude that it is raining (as it is much more likely that the sprinklers have been on; Pearl, 1988). Whenever some facts are strongly established at the outset of the reasoning process (e.g., the defendant's character in Experiment 3), these will tend to function as premises from which conclusions are derived. Our claim is simply that natural human decision making does not exclusively depend on strictly directional inferences.

Other evidence supports the possibility that bidirectional, coherence-based reasoning may be as basic to human inference as is directed deduction. The difficulty of grasping the distinction between hypotheses and evidence (Kuhn, 1989; Ranney et al., 1996) is consistent with the primacy of coherence-based reasoning, which has a bidirectional component. There is evidence that teaching formal rules for reasoning can be facilitated by grounding the training on natural reasoning strategies (Cheng, Holyoak, Nisbett, & Oliver, 1986); accordingly, in teaching reasoning it may prove useful to take into account a human predilection for making coherence-based inferences. We suspect that constraint satisfaction based on bidirectional inferences plays an important role not only in everyday decision making by ordinary people, but also in professional decision-making activities, such as judicial reasoning (Simon, in press), politics, and medicine.

In the domain of politics, the generation of patterns of coherent beliefs through the creation of constraint networks

may help explain the emergence of political ideologies. Lakoff (1996) observed that ideologies may consist of clusters of beliefs and attitudes that appear to lack logical links. For example, current American conservatism often combines favoring free access to guns with favoring restricted access to abortion, whereas liberalism tends to combine the opposite views; however, it is unclear what the two points of dispute have to do with one another. Lakoff proposed that conservatism and liberalism are respectively grounded in two alternative analogical extensions of the family. This possibility is certainly compatible with the present study in that an analogy can provide a potent basis for a set of interrelated inferences. However, coherence-based reasoning provides a yet broader conception of how coalitions of beliefs and attitudes can arise. An analogy may provide the core, but, more generally, variables that cohere with favorably viewed outcomes will tend to come to cohere with each other. In Experiment 3, for example, a favored analogy appeared to tip the balance toward a certain decision in the Infoscience case; in doing so, it carried with it a certain assessment of who was responsible for the company's profits. The latter inference was not directly based on the analogy but, rather, cohered with a decision that in turn cohered with the favored analogy. Such indirect connections may foster the emergence of loosely connected attitudes that cohere to form a stable cluster.

More generally, the present study supports the view that many high-level cognitive processes, including inference, belief and attitude change, decision making, and evaluation of competing explanations, can be understood as interacting aspects of an overarching process of soft constraint satisfaction. When viewed from this perspective, human understanding appears less like a serial process of logical deduction, moving rigidly from established premises to new conclusions, and more like the solution to a complex puzzle in which the individual pieces must be reorganized and transformed to form a coherent whole.

References

Abelson, R. P., & Rosenberg, M. (1958). Symbolic psycho-logic: A model of attitudinal cognition. *Behavioral Science, 3*, 1–8.

Bem, D. J., & McConnell, H. K. (1970). Testing the self-perception explanation of dissonance phenomena: On the salience of premanipulation attitudes. *Journal of Personality and Social Psychology, 14*, 23–31.

Bentler, P. M. (1995). *EQS structural equations program manual.* Encino, CA: Multivariate Software.

Blanchette, I., & Dunbar, K. (1997). Constraints underlying analogy use in a real-world-context: Politics. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the nineteenth annual conference of the Cognitive Science Society* (p. 867). Hillsdale, NJ: Erlbaum.

Cheng, P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18*, 293–328.

Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences.* Hove, England: Erlbaum.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Festinger, L. (Ed.). (1964). *Conflict, decision, and dissonance.* Stanford, CA: Stanford University Press.

Fischhoff, B. (1982). For those condemned to study the past: Heuristics and biases in hindsight. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 335–351). New York: Cambridge University Press.

Goethals, G. R., & Reckman, R. F. (1973). The perception of consistency in attitudes. *Journal of Experimental Social Psychology, 9*, 491–501.

Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology, 21*, 107–111.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Higgins, E. T., & Liberman, A. (1994). Memory errors from a change of standard: A lack of awareness or of understanding? *Cognitive Psychology, 27*, 227–258.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13*, 295–355.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought.* Cambridge, MA: MIT Press.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Khong, Y. F. (1992). *Analogies at war.* Princeton, NJ: Princeton University Press.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review, 96*, 674–689.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel constraint satisfaction theory. *Psychological Review, 103*, 284–308.

Lakoff, G. (1996). *Moral politics.* Chicago: University of Chicago Press.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*, 585–589.

Markman, A. B., & Medin, D. L. (1995). Similarity and alignment in choice. *Organizational Behavior and Human Decision Processes, 63*, 117–130.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1: An account of basic findings. *Psychological Review, 88*, 375–407.

McGuire, W. J. (1960). Cognitive consistency and attitude change. *Journal of Abnormal and Social Psychology, 60*, 345–353.

McGuire, W. J. (1981). The probabilogical model of cognitive structure and attitude change. In R. E. Petty, T. M. Ostrom, & T. C. Brock (Eds.), *Cognitive responses in persuasion* (pp. 291–307). Hillsdale, NJ: Erlbaum.

McGuire, W. J. (1990). Dynamic operations of thought systems. *American Psychologist, 45*, 504–512.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference.* San Mateo, CA: Morgan Kaufmann.

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. (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.

Ranney, M., Schank, P., Hoadley, C., & Neff, J. (1996). "I know one when I see one": How (much) do hypotheses differ from evidence? In R. Fidel, B. H. Kwasnik, C. Beghtol, & P. J. Smith (Eds.), *Advances in classification research, Vol. 5* (American Society for Information Science Monograph Series, pp. 141–158). Medford, NJ: Learned Information.

Ranney, M., & Thagard, P. (1988). Explanatory coherence and belief revision in naive physics. In *Proceedings of the tenth annual conference of the Cognitive Science Society* (pp. 426–432). Hillsdale, NJ: Erlbaum.

Read, S. J., & Miller, L. C. (1993). Rapist or "regular guy": Explanatory coherence in the construction of mental models of others. *Personality and Social Psychology Bulletin, 19,* 526–541.

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.), *Belief, reasoning, and decision-making: Psycho-logic in honor of Bob Abelson* (pp. 209–235). Hillsdale, NJ: Erlbaum.

Ross, M., & Shulman, R. F. (1973). Increasing the salience of initial attitudes: Dissonance versus self-perception theory. *Journal of Personality and Social Psychology, 28,* 138–144.

Schank, P., & Ranney, M. (1991). The psychological fidelity of ECHO: Modeling an experimental study of explanatory coherence. In J. K. Kruschke (Ed.), *Proceedings of the thirteenth annual conference of the Cognitive Science Society* (pp. 599–604). Hillsdale, NJ: Erlbaum.

Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. I and II. *Psychometrika, 27,* 125–140.

Shultz, T. R., & Lepper M. R. (1996). Cognitive dissonance reduction as constraint satisfaction. *Psychological Review, 103,* 219–240.

Simon D. (in press). Towards a psychology of judicial reasoning. *Rutgers Law Journal.*

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., & Holyoak, K. J. (1996). Pragmatics in analogical mapping. *Cognitive Psychology, 31,* 307–346.

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.

Taplin, J. E., & Staudenmayer, H. (1973). Interpretation of abstract conditional sentences in deductive reasoning. *Journal of Verbal Learning and Verbal Behavior, 12,* 530–542.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences, 12,* 435–467.

Thagard, P. (1992). *Conceptual revolutions.* Princeton, NJ: Princeton University Press.

Thagard, P., Holyoak, K. J., Nelson, G., & Gochfeld, D. (1990). Analog retrieval by constraint satisfaction. *Artificial Intelligence, 46,* 259–310.

Thagard, P., & Millgram, E. (1995). Inference to the best plan: A coherence theory of decision. In A. Ram & D. B. Leake (Eds.), *Goal-driven learning* (pp. 439–454). Cambridge, MA: MIT Press.

Ullman, J. B. (1996). Structural equation modeling. In B. G. Tabachnick & L. S. Fidell (Eds.), *Using multivariate statistics* (3rd ed., pp. 709–812). New York: HarperCollins.

Wharton, C. M., Holyoak, K. J., Downing, P. E., Lange, T. E., Wickens, T. D., & Melz, E. R. (1994). Below the surface: Analogical similarity and retrieval competition in reminding. *Cognitive Psychology, 26,* 64–101.

Wharton, C. M., Holyoak, K. J., & Lange, T. E. (1996). Remote analogical reminding. *Memory & Cognition, 24,* 629–643.

Wixon, D. R., & Laird, J. D. (1976). Awareness and attitude change in the forced-compliance paradigm: The importance of when. *Journal of Personality and Social Psychology, 34,* 376–384.

Young, F. W., & Harris, D. F. (1990). Multidimensional scaling: Procedure ALSCAL. In *SPSS base system user's guide* (pp. 396–461). Chicago: SPSS Inc.

Appendix A

Pretest Used in Experiment 1

I. Instructions

Please give your opinion about the following assortment of issues. Each issue is preceded by a brief summary of relevant information. You will then be asked to rate the extent to which you agree or disagree with a statement about the issue. The issues concern factual situations, public policy, business situations, and legal affairs. You are not expected to have any expert knowledge. You should use common sense in making your ratings. That is, you should apply your sense of reasonableness, fairness, and good policy. The issues are unrelated, so simply consider each issue independently. You might find that the information contained in the statements is less than you would like to have; nonetheless, respond as best you can based on the information provided.

For each of the statements about issues, please rate the extent to which you agree or disagree. Answer all questions using the following scale

Disagree										Agree
−5	−4	−3	−2	−1	0	1	2	3	4	5
strongly	moderately	weakly			neutral		weakly	moderately		strongly

II. Questions Related to Each of Six Points of Dispute

[Please note that an asterisk indicates questions for which the sign on ratings was reversed in computing Q-scores, that is, those questions for which a positive rating signified agreement with Smith's position.]

Truth

1. A software development company is gradually losing its competitive edge and has not upgraded its original products. Its managers are reported to be quarreling among themselves. Sales

are dropping. The company is constantly losing money and borrowing more to fund its operations, while its stock price has been slipping steadily for months.

An observer believes that the company is in a mess and that its situation is hopeless. The observer's belief is well-founded.

Cause

2. A high-tech company succeeded remarkably with a popular line of products. Its success, however, drew a great deal of competition, and the profitability in that field began to shrink. Rather than compete in that fierce market, the company made a strategic decision to direct its efforts toward new markets. It devoted all its resources to the development of an innovative line of products, at the expense of updating its original ones. During this secret development phase, sales dropped and additional funds were borrowed to fund the development.

One month before the surprise launching of the new products, an observer spread a message stating that the company is in a mess and the situation is hopeless. Days after the message was distributed, the company's stock plummeted sharply, causing its creditors to hustle for the money owed to them. The company's available assets were inadequate to meet the rush; two weeks later it filed for bankruptcy.

a. The message caused the company's collapse.

*b. The company collapsed because of mismanagement.

Motive

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

a. The investor's action was motivated primarily by vindictiveness; he was a sore loser who was happy to make money from the company's success but unwilling to sit through tough times.

*b. The investor's main intention was to prevent other innocent potential investors from being misled into a bad investment.

Regulation

4. A company's regulation states, *"Prior to taking any action that is reasonably expected to cause serious harm to the company, an investor must first notify the managers, unless he or she has good reason to believe that management will not respond to the notification."* An executive had previously denied allegations of problems in the company. Without notifying management, an investor spreads a message that the company is in a mess and its situation is hopeless.

The investor's message violated the company's regulation.

Speech

5. The Internet enables people to spread information rapidly to very large audiences. Society will be better off if the law prohibits abuses of the system, such as transmission of libelous statements that damage the reputation of individuals or companies.

As a matter of policy, communications over the Internet ought to be regulated by law.

*6. The Internet is becoming the new "Marketplace of Ideas."

As a matter of policy, keeping the Internet open to the free exchange of viewpoints and information is a vital social need.

Analogy

7. A newspaper is defined as: "any publication intended for the distribution and dissemination of news, facts, or opinions to broad audiences." An electronic "bulletin board" is a forum on computer networks (such as the Internet) where subscribers post messages that are open to other subscribers.

As a matter of legal policy, messages posted on electronic bulletin boards should be treated like items published in newspapers.

*8. A telephone system is defined as "a network of interconnected lines used to transmit and receive voice or data from one extension to one or more other extensions." An electronic "bulletin board" is a forum on computer networks (such as the Internet) where subscribers post messages that are open to other subscribers.

As a matter of legal policy, messages posted on electronic bulletin boards should be treated like messages sent over a telephone network.

Appendix B

Materials for Case *Caught in the Net*

I. General Instructions [2-Phase and 3-Phase Conditions]

In this experiment you will be asked to play the role of a judge on the Federal Court of Appeals. You have been assigned to decide a libel suit. A company is suing one of its investors for a message he posted on an electronic "bulletin board" located on the Internet. The case is unique in that it is one of the first legal disputes to have been spawned by the recent emergence of electronic communication technologies.

In deciding this case, you will try to justify your image as an up-and-coming wise judge. You might see strengths and weaknesses in both positions, but it is your duty to decide in favor of one party or the other—that is, to announce which side wins. The decision is yours. You should base it as soundly and fairly as possible on the available facts and arguments from both sides.

You will later notice that each party cites a precedent in support of its case. (Precedents are previously decided cases that have some relevant similarities to the current case.) You are not allowed to change precedents, but it is up to you to determine the extent to

which any precedent is pertinent to the current case. In general, this experiment does not expect you to have any legal knowledge. You should use common sense in deriving your decision. That is, you should apply your sense of reasonableness, fairness, and good policy.

[The following section was added for the 3-phase condition only.]

As you sit down to study the case, you hear that a very similar case is about to be decided by Judge Brown, sitting on another Federal Appeals Court. If the two decisions (yours and Judge Brown's) turn out contradictory to each other, it will cause significant confusion as to what the state of the law is. Furthermore, Judge Brown has an outstanding reputation as a careful and thoughtful judge, so reaching a result that contradicts that of Judge Brown would be embarrassing for you.

Because of Judge Brown's impending decision, you have devised the following plan. You will begin to work on your case as you would any other. That is, you will carefully study the materials and evaluate all the facts and arguments involved. However, you will avoid reaching a final decision until Judge Brown's verdict is returned. You will then consider your evaluation of the current case in light of Judge Brown's decision and decide upon your verdict.

You will first see the information about your case. After you have had time to consider it carefully, you will be asked for to evaluate the facts and arguments. Then you will be told Judge Brown's decision in the related case.

II. Factual Summary

On this page you will read a summary of the basic facts of the case. These facts are not in dispute, as both parties have testified to their accuracy. Read this summary carefully and be sure you understand these basic facts. Take as much time as you feel you need. You will be able to look back at this page later if you need to review the material. When you are done reading this factual summary, turn to the next page.

Summary of the Facts of Quest v. Smith

A software company by the name of Quest is suing Jack Smith for harm he allegedly caused it by a derogatory statement he made about it. The statement was made on Cash-Chat, an electronic bulletin board devoted to discussion about financial investments. Cash-Chat is intended primarily for Wall Street brokers, though small investors have access to it too.

Quest's financial situation had been running a volatile course for some time. Like many other high-powered software developers of its kind, it emerged swiftly from anonymity and captured an eminent position in its market. The company's rise brought handsome profit to its initial investors. Over the last two years, however, Quest had been giving its investors nothing but aggravation. The company was gradually losing its competitive edge. *Business Week* reported that the company was suffering from "excessive agitation and animosity at the management level." The last three financial reports showed increasing losses, forcing the company to borrow more money to fund its operations. Quest's stock price had been slipping steadily for several months.

Jack Smith owned some shares in Quest. At first he made a profit from his investment in the company, but more recently he has been watching his investment shrivel. He discovered that the company had not been upgrading its products. Smith was becoming increasingly frustrated with the company's performance; he was especially angered by its top executive's brazen denial of any troubles and hyperbolic promises for a great future. The next day, Smith placed the following message on Cash-Chat: "My research indicates that the company is in debt up to its neck. No cash, no sales, no profit, and worst of all, no product upgrading. Managers are capable of nothing but internal brawling. The company is a sham. The situation is hopeless!"

Days after Smith's message, the company's stock plummeted a sharp 35%, causing its creditors to make immediate demands of Quest to repay the money the company owed to them. The company's available assets were inadequate to meet the rush; two weeks later it filed for bankruptcy.

The company then sued Smith for libel. It subsequently became known that at the time of the collapse, Quest was just one month away from launching a new line of software packages that it had been developing secretly. Analysts say that the new line appeared to be quite promising.

III. Summary of Opposing Arguments

On this page you will read summaries of the arguments offered by each of the two parties. You should consider all of these arguments very carefully. Take as much time as you feel you need. You will be able to look back at this page later if you need to review the material. When you are done reading these arguments, turn to the next page.

Quest's Arguments

[Version in which Quest's arguments preceded Smith's:]

Quest claims that Smith's message was unfounded. The company asserts that Smith's false message caused the company's collapse by frightening the company's investors and creditors. According to Quest, Smith's action was an act of vindictiveness; he was a sore loser who was happy to make money from Quest's initial success but was unwilling to sit through tough times.

Quest had a company regulation: *"Prior to taking any action that is reasonably expected to cause serious harm to the company, an investor must first notify the managers, unless he or she has good reason to believe that management will not respond to the notification."* Quest argues that Smith's message violated this regulation. Quest asserts that Smith should have known his message would seriously harm the company and was therefore required to give adequate notice to management, who would have certainly rebutted his allegations.

Quest argues that as a matter of policy, communications over the Internet ought to be regulated by law. The Internet enables people to spread information rapidly to very large audiences. According to Quest, Society will be better off if the law prohibits abuses of the system, such as transmission of libelous statements that damage the reputation of individuals or companies.

The company claims that messages posted on electronic bulletin boards should be treated like items published in newspapers. It cites a judicial precedent stating that defamatory expressions printed in newspapers can give rise to a claim of libel. That precedent defined a newspaper as: *"any publication intended for the distribution and dissemination of news, facts, or opinions to broad audiences."*

Smith's Arguments

Smith insists his message was well-founded and accurate. He argues that the company collapsed because of its poor performance, rather than as a result of his message. Smith asserts that his main intention was to prevent other innocent investors from being misled into a bad investment.

Smith argues his action was consistent with the company's regulation (see above). Smith insists that his message couldn't reasonably have been expected to harm the company in any way. The company's top executive had already denied that the company was in trouble, giving him good reason to believe management

would not respond to notification. According to Smith, delaying his message any longer would have caused many other innocent potential investors to be harmed by Quest's imminent collapse.

Smith argues that the Internet is becoming the new "Marketplace of Ideas." He argues that as a matter of policy, keeping the Internet open to the free exchange of viewpoints and information is a vital social need.

Smith contends that messages posted on the Internet should be treated like messages made over a telephone system. He cites a judicial precedent stating that defamatory messages transmitted over the telephone do not amount to libel. That precedent defined a telephone system as "*a network of interconnected lines used to transmit and receive voice or data from one extension to one or more other extensions.*"

IV. Instructions for Interim Test [3-Phase Condition Only]

At this stage you are still waiting to hear Judge Brown's verdict and are suspending your own decision. However, you may have a preliminary leaning toward one of the two parties. Please give your preliminary leaning and evaluations below. You may look back at the summaries of facts and arguments as much as you like. You will still be free to give any verdict you think is appropriate later.

V. Decision Instructions

Instructions For Making Your Decision

[The following paragraph was added for the 3-phase condition only:]

You now hear that Judge Brown has become ill and will not be deciding the similar case for several more weeks. As you cannot delay your decision any longer, you must now go ahead and make your own decision based on your independent judgment. So, regardless of your earlier tentative leaning, you must now reach the best possible verdict.

[The following section was provided to both the 2-phase and 3-phase conditions:]

In deciding this case, you will try to justify your image as an up-and-coming wise judge. As stated earlier, you might see strengths and weaknesses in both positions, but it is your duty to decide in favor of one party or the other—that is, to announce which side wins. The decision is yours. You should base it as soundly and fairly as possible on the available facts and arguments from both sides.

Each party cited a precedent in support of its case. You are not allowed to change precedents, but it is up to you to determine the extent to which any precedent is pertinent to the current case. In general, this experiment does not expect you to have any legal knowledge. You should use common sense in deriving your decision. That is, you should apply your sense of reasonableness, fairness, and good policy.

Fortunately for you, in cases of this sort the only thing you have to announce publicly is your verdict. You are not obliged to write a legal opinion. However, as you reach your decision you must do your best to ensure that it corresponds to your evaluation of the facts and arguments. You will also be asked a series of questions regarding other aspects of the case.

[On the next page, participants indicated their verdict and rated their confidence on a 5-point scale (posttest).]

VI. Questions Related to Each of Six Points of Dispute

[Please note that an asterisk indicates questions for which the sign on ratings was reversed in computing Q-scores, that is, those questions for which positive rating signified agreement with Smith's position.]

For each of the following arguments that were made by Quest or Jack Smith, please rate the extent to which you agree or disagree. Answer all questions using the following scale:

Disagree										Agree
−5	−4	−3	−2	−1	0	1	2	3	4	5
strongly	moderately	weakly		neutral		weakly		moderately		strongly

Truth

1. Smith's message was well-founded.

Cause

2a. Smith's message caused Quest's collapse.

*2b. Quest collapsed because of poor performance.

Motive

3a. Smith's action was motivated primarily by vindictiveness.

*3b. Smith's main intention was to prevent other innocent potential investors from being misled into a bad investment.

Regulation

4. Smith's message violated the company's regulation concerning prior notification to management.

Speech

5. As a matter of policy, communications over the Internet ought to be regulated by law.

*6. As a matter of policy, keeping the Internet open to the free exchange of viewpoints and information is a vital social need.

Analogy

7. Messages posted on electronic bulletin boards should be treated like items published in newspapers (see definition in *Quest's* argument), and thus can give rise to a claim of libel.

*8. Messages posted on electronic bulletin boards should be treated like messages sent over a telephone network (see definition in *Smith's* argument), and thus cannot give rise to a claim of libel.

Appendix C

Manipulation of Defendant's History and Character in Summary of Facts for Quest Case (Experiment 3)

I. Good Smith Version

Jack Smith is an investor who has an impressive record of careful scrutiny of companies that sell their stock to the public. He has often been successful in bringing about major management changes that have greatly benefited individual shareholders. A software company by the name of Quest is now suing Smith for harm he allegedly caused it by a statement he made about the company. The statement was made on an electronic bulletin board located on the Internet, which serves as a forum for discussion about financial investments.

Smith pays careful attention to the inner workings of the companies in which he invests. He never hesitates to share information that he feels will protect other shareholders, even if it sometimes means he takes a personal loss as a consequence. He has a reputation as a scrupulously honest watchdog, who has sometimes been compared to Ralph Nader. His method has been as follows: He first invests small amounts in a company and obtains information about its operations. After discovering the company's weak spots, Smith is a leader in making the concerns of the shareholders known to the executives of the company, often offering specific suggestions on how the company's problems can be solved. When the management is responsive and implements Smith's suggestions for change, the results have often been a spectacular turnaround. By his actions Smith has often caused both the company and all other investors to increase their profits.

Quest's financial situation had been running a volatile course for some time. Like many other high-powered software developers of its kind, it emerged swiftly from anonymity and captured an eminent position in its market. The company's rise brought handsome profit to its initial investors. Over the last two years, however, Quest had been giving its investors nothing but aggravation. The company was gradually losing its competitive edge. *Business Week* reported that the company was suffering from "excessive agitation and animosity at the management level." The last three financial reports showed increasing losses, forcing the company to borrow more money to fund its operations. Quest's stock price had been slipping steadily for several months.

After investing in Quest, Smith discovered that the company was failing to upgrade its current products. Smith was becoming increasingly concerned with the company's performance; he was especially alarmed by its top executive's brazen denial of any troubles and hyperbolic promises for a great future. One day after Smith heard the executive denying any troubles in Quest, he placed the following message on the bulletin board: "My research indicates that the company is in debt up to its neck. No cash, no sales, no profit, and worst of all, no product upgrading. Managers are capable of nothing but internal brawling. The company is a sham. The situation is hopeless!"

Days after Smith's message, the company's stock plummeted a sharp 35%, causing its creditors to make immediate demands of Quest to repay the money the company owed to them. The company's available assets were inadequate to meet the rush; two weeks later it filed for bankruptcy.

Quest is now suing Smith for libel. It subsequently became known that at the time of the collapse, Quest was just one month away from launching a new line of software packages that it had been developing secretly. Analysts say that the new line appeared to be quite promising.

II. Bad Smith Version

Jack Smith is an investor who has a shocking record of financial wrongdoing. He has already been convicted three times for deliberately ruining companies in order to buy their assets for less than their true value. A software company by the name of Quest is now suing Smith for harm he allegedly caused it by a statement he made about the company. The statement was made on an electronic bulletin board located on the Internet, which serves as a forum for discussion about financial investments.

On at least three previous occasions, Smith has ruined good companies and then bought their assets at rock-bottom prices. His method has been as follows: He first invests small amounts in a company and obtains information about its operations. After discovering the company's weak spots, he spread libelous information that was sure to cause serious harm to the company. Smith's messages cause the company's value to plunge, which in turn causes the company's creditors to demand the money owed to them. Following these events, Smith buys whatever was left of the company for almost free. He later sells the remains of the company at a profit. By his actions Smith has often caused all other investors to lose their entire investments.

Quest's financial situation had been running a volatile course for some time. Like many other high-powered software developers of its kind, it emerged swiftly from anonymity and captured an eminent position in its market. The company's rise brought handsome profit to its initial investors. Over the last two years, however, Quest had been giving its investors nothing but aggravation. The company was gradually losing its competitive edge. *Business Week* reported that the company was suffering from "excessive agitation and animosity at the management level." The last three financial reports showed increasing losses, forcing the company to borrow more money to fund its operations. Quest's stock price had been slipping steadily for several months.

After investing in Quest, Smith discovered that the company was secretly developing a new line of innovative software products, and that because of the intensive development effort, it had refrained from upgrading the current products. One day after Smith heard a top executive denying any troubles in Quest, Smith placed the following message on the bulletin board: "My research indicates that the company is in debt up to its neck. No cash, no sales, no profit, and worst of all, no product upgrading. Managers are capable of nothing but internal brawling. The company is a sham. The situation is hopeless!"

Days after Smith's message, the company's stock plummeted a sharp 35%, causing its creditors to make immediate demands of Quest to repay the money the company owed to them. The company's available assets were inadequate to meet the rush; two weeks later it filed for bankruptcy. Within days after the crash, one of Smith's investment companies offered to buy out Quest's products for an extremely low price.

Quest is now suing Smith for libel. It subsequently became known that at the time of the collapse, Quest was just one month away from launching the new line of software packages that it had been developing secretly. Analysts say that the new line appeared to be quite promising.

(Appendixes continue)

Appendix D

Materials for Case *The Bonus Dispute at Infoscience* (Experiment 3)

I. General Instructions

In your role as a judge on the Federal Court of Appeals, you have now been assigned to decide a labor–management dispute. The parties request that you clarify the contract between a company and its employees. Unfortunately for you, you may not suggest a compromise; you must decide the case in favor of either one party or the other. Once again you should use common sense in deriving your decision. That is, you should apply your sense of reasonableness, fairness, and good policy.

On this page you will read a summary of the basic facts of the case. These facts are not in dispute, as both parties have testified to their accuracy. Read this summary carefully and be sure you understand these basic facts. Take as much time as you feel you need. You will be able to look back at this page later if you need to review the material. When you are done reading this factual summary, turn to the next page.

II. Factual Summary

Summary of the Facts of Infoscience v. Employees

Infoscience is an electronic bulletin board operating out of a suburb of Indianapolis. Infoscience concentrates on the natural sciences and is one of the most prominent bulletin boards used in scientific research. The service serves some 8,000 scientists, universities, and research labs across the country. The service has grown steadily since its inception in 1984, and it has been profitable throughout the past six years. In recent months, however, Infoscience has been troubled by a dispute between its management and its labor force of employees. The dispute concerns the bonus employees are to be paid for last year's accomplishment. Infoscience's owner and general manager, Mike Johnson, has proposed a bonus of 6%, whereas the employees maintain that they deserve a bonus of 14%.

Both sides are basing their positions on the employment contract they signed one year earlier. In the contract, the parties agreed that Infoscience would pay an annual bonus based on two factors. First, the bonus was to be similar to that paid by other information-service companies located in Indianapolis. Second, the bonus should be based on an evaluation of the employees' overall contribution to the company's profitability.

The problem arose when Mike Johnson and the employees failed to agree on how to interpret the two factors specified by the contract. They disagreed about precisely what kind of company provided the appropriate comparison for Infoscience (a point the contract had not spelled out) and also about the nature of the employees' contribution to the company's profitability over the past year. Both sides expressed an interest in establishing guidelines for their future relations, so they decided by mutual consent to bring the dispute to court without delay.

III. Summary of Opposing Arguments

Infoscience's Arguments

[The following is the version in which Infoscience cites newspaper analogy and Infoscience's arguments precede the employees' arguments:]

Mike Johnson, the president of Infoscience, presents the company's case. He states that of the handful of information-service companies located in Indianapolis, the one that is most similar to a bulletin board is the city's newspaper, *The Indi Star*. He cites a judicial precedent defining a newspaper as: "*any publication intended for the distribution and dissemination of news, facts, or opinions to broad audiences.*" Johnson argues that an electronic bulletin board clearly serves the function of a newspaper; accordingly, the bonus at the local newspaper is the relevant comparison for determining the bonus for Infoscience employees. The *Indi Star* has just given its employees bonuses of 5%; hence, Johnson argues that 6% is a generous bonus for Infoscience employees.

Johnson further claims that the employees deserve only a small part of the credit for the company's profits. He asserts that the growth in the bulletin board's capacity over the past year was more strongly related to the advanced computer system that was installed during the past year than to the workers' contribution.

The Employees' Arguments

The employees claim that of the information-service companies located in Indianapolis, the one which most resembles the bulletin board is the county's telephone company. They cite a judicial precedent defining a telephone system as: "*a network of interconnected lines used to transmit and receive voice or data from one extension to one or more other extensions.*" The employees argue that an electronic bulletin board clearly constitutes a network of the same sort as a telephone system; accordingly, the bonus at the local telephone company is the relevant comparison for determining the bonus for Infoscience employees. For last year's performance, the telephone company gave its employees a bonus of 15%; hence the Infoscience employees argue that 14% is a conservative bonus for them.

The employees state, further, that they deserve full credit for the bulletin board's profitability. They claim that learning to operate the complex new computer system and having to cope with its never-ending bugs demanded an unusual effort on their behalf. They emphasize that they are working more hours with the new computer than they previously did.

IV. Posttest Questions Related to Each of Two Points of Dispute

[Please note that the sign of rating for Q-scores was counterbalanced across participants.]

Analogy

1. For the purpose of determining the bonus at Infoscience, the bulletin board should be treated like the Indianapolis newspaper.
2. For the purpose of determining the bonus at Infoscience, the bulletin board should be treated like the Indianapolis telephone company.

Credit

3. Infoscience's profitability in the preceding year was due primarily to the new computer system.
4. Infoscience's profitability in the preceding year was due primarily to the employees' hard work.

V. Pretest Questions Related to the Credit Factor

3'. A company that operates a bulletin board installs a new computer system. The following year it makes a substantial profit. An observer infers that the profitability was due to the newly installed computer system.

The observer's inference is correct.

4'. A company that operates a bulletin board installs a new computer system. The following year it makes a substantial profit. An observer reports that while productivity rose, having to learn the complex system and to cope with its bugs demanded an unusual effort on behalf of the company's employees. The observer infers that the increase in profitability was due to the employees' efforts.

The observer's inference is correct.

Received March 28, 1997
Revision received October 23, 1997
Accepted February 4, 1998 ■