

Explanation-Based Decision Making

Nancy Pennington and Reid Hastie

University of Chicago Northwestern University

Abstract

In complex decision tasks the decision maker frequently constructs a summary representation of the relevant evidence in the form of a causal explanation and relies on that representation, rather than the "raw" evidence base, to select a course of action from a choice set of decision alternatives. We introduce a general model for this form of decision making, called explanation-based decision making, because of the central role played by the intervening evidence summary. Several original empirical studies of judicial decision making, a prototype of the class of explanation-based decision tasks, are reviewed and the findings are adduced in support of the explanation-based decision model. In legal decision making tasks subjects spontaneously construct evidence summaries in the form of stories comprising the perceived underlying causal relationships among decision relevant events. These explanations are primary mediators (i.e., causes) of the subjects' decisions.

Many important decisions in engineering, medical, legal, policy, and diplomatic domains are made under conditions where a large base of implication-rich, conditionally-dependent pieces of evidence must be evaluated as a preliminary to choosing an alternative from a set of prospective courses of action. We propose that a general model of explanation-based decision making describes behavior under these conditions (Pennington & Hastie, 1986). According to the explanation-based decision model, decision makers begin their decision process by constructing a causal model to explain the available facts. Concomitant with or subsequent to the construction of a causal model of the evidence, the decision maker is engaged in a separate activity to learn or create a set of alternatives from which an action will be chosen. A decision is made when the causal model of the evidence is successfully matched to an alternative in the choice set. The three processing stages in the explanation-based decision model are shown in Figure 1.

The distinctive assumption in our explanation-based approach to decision making is the hypothesis that decision makers construct an intermediate summary representation of the evidence and that this representation, rather than the original "raw" evidence, is the basis of the final decision. Interposition of this organization facilitates evidence comprehension, directs inferencing, enables the decision maker to reach a decision, and determines the confidence assigned to the accuracy or success of the decision. This means that the locus of theoretical accounts for differences in decisions rendered by different individuals, systematic biases shared by many individuals, and the effects of most variations in decision task characteristics will usually lie in the evidence evaluation stage of the decision process.

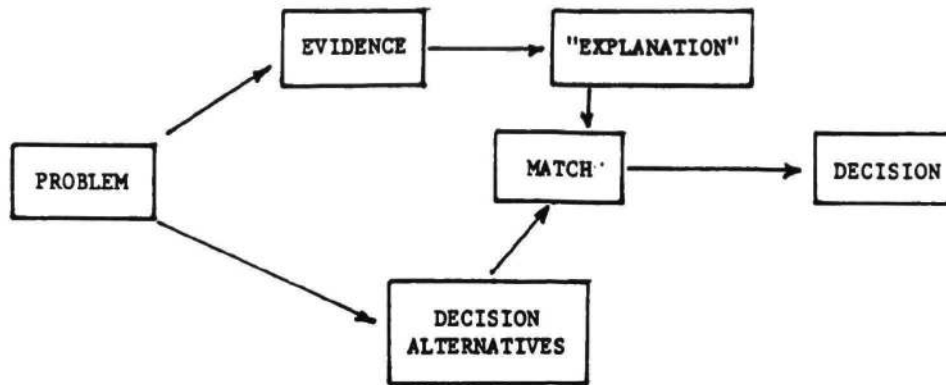


Figure 1

Overview of the Processing Stages of the Explanation-based Model

The structure of the causal model constructed to explain the evidence will be specific to the decision domain. For example, we have proposed that a juror uses narrative story structures to organize and interpret evidence in criminal trials. Different causal rules and structures will underlie an internist's causal model of a patient's condition and its precedents (Pople, 1982), an engineer's mental model of an electrical circuit (de Kleer & Brown, 1983), a merchant's image of the economic factors in a resort town (Hogarth, Michaud, & Mery, 1980), or a diplomat's causal map of the political forces in the Middle East (Axelrod, 1976). Thus, a primary task in research on explanation-based decision making is the identification of the type of intermediate summary structure that is imposed on* evidence by decision makers in a specific domain of decision making. This is in contrast with earlier process-oriented calculational models where the theoretical focus was on attentional processes and the computations whereby separate sources of information were integrated into a unitary value or utility (Anderson, 1981; Edwards, 1954; Kahneman & Tversky, 1979).

Explanation-based Decision Making in Judicial Decisions

In the present paper we concentrate on the example of juror decision making. The juror's decision task is a prototype of the tasks to which the explanation-based model should apply: First, a massive "database" of evidence is input at trial, frequently requiring several days to present. Second, the evidence comes in a scrambled sequence, usually several witnesses and exhibits convey pieces of a historical puzzle in a jumbled temporal sequence. Third, the evidence is piecemeal and gappy in its depiction of the historical events that are the focus of reconstruction: event descriptions are incomplete, usually some critical events were not observed by the available witnesses, and information about personal reactions and motivations is not presented (often because of the rules of evidence). Finally, subparts of the evidence (e.g., individual sentences or statements) are interdependent in their probative implications for the verdict. The meaning of once statement cannot be assessed in isolation because it depends on the meanings of several related statements.

Evidence Summary. Empirical research has demonstrated that the juror's "explanation" of legal evidence takes the form of a "story" in which causal and intentional relations among events are prominent (Bennett & Feldman, 1981; Hutchins, 1980; Pennington, 1981; Pennington & Hastie, 1986). The story is constructed from information explicitly presented at trial and knowledge possessed by the juror. Two kinds of knowledge are critical: (a) expectations about what makes a complete story and (b) knowledge about

events similar in content to those that are the topic of dispute.

General knowledge about the structure of human purposive action sequences, characterized as an episode schema, serves to organize events according to the causal and intentional relations among them as perceived by the juror. An episode schema specifies that a story should contain initiating events, goals, actions, consequences, and accompanying states, in a particular causal configuration (Mandler, 1980; Pennington & Hastie, 1986; Rumelhart, 1977; Stein & Glenn, 1979; Trabasso & van den Broek, 1985). Each component of an episode may consist of an episode so that the story the juror constructs can be represented as a hierarchy of embedded episodes. The highest level episode characterizes the most important features of "what happened." Components of the highest level episode are elaborated in terms of more detailed event sequences in which causal and intentional relations among subordinate story events are represented. Expectations about the kinds of information necessary to make a story tell the juror when important pieces of the explanation structure are missing and when inferences must be made. Knowledge about the structure of stories allows the juror to form an opinion concerning the completeness of the evidence, the extent to which a story has all its parts.

More than one story may be constructed by the juror, however one story will usually be accepted as more coherent than the others. Coherence combines judgments of completeness, consistency, and plausibility. Consistency concerns the extent to which the story does not contain contradictions and the plausibility of alternative stories may be assessed by comparing story sequences to known or imagined events in the real world. If more than one story is judged to be coherent, then the story will lack uniqueness and great uncertainty will result. If there is one coherent story, this story will be accepted as the explanation of the evidence and will be instrumental in reaching a decision.

Choice Set. The decision maker's second major task is to learn or to create a set of potential solutions or action alternatives that constitute the choice set. In some decision tasks the potential actions are given to the decision maker (instructions from the trial judge on verdict alternatives) or known beforehand (treatment options available to a physician). In others, creation of alternatives is a major activity of the decision maker (for example, drafting alternate regulations for industrial waste disposal, planning alternate marketing strategies, or negotiating alternate acceptable trade contracts). These solution design tasks may invoke their own (embedded) decision tasks.

In criminal trials the information for this processing stage is given to jurors at the end of the trial in the judge's instructions on the law. The process of learning the verdict categories is a one-trial learning task in which the material to be learned is very abstract. Interference may occur from jurors' prior knowledge of concepts such as first degree murder, manslaughter, armed robbery, etc. The juror attempts to learn the defining features (elements of the crime) of each verdict alternative and a decision rule specifying their appropriate combination. We hypothesize that the conceptual unit is a category (frame) defined by a list of criterial features referring to identity, mental state, circumstances, and actions linked conjunctively or disjunctively to the verdict alternative (Kaplan, 1978; Pennington & Hastie, 1981).

Match Process. The final stage in the global decision process involves matching solution alternatives to the summary evidence representation to find the most successful pairing. Confidence in the final decision will be

partly determined by the goodness-of-fit of the evidence-solution pairing selected and the uniqueness of the winning combination when compared to alternative pairings. Because verdict categories are unfamiliar concepts, the classification of a story into an appropriate verdict category is likely to be a deliberate process. For example, a juror may have to decide whether a circumstance in the story such as "pinned against a wall" constitutes a good match to a required circumstance, "unable to escape," for a verdict of Not Guilty by Reason of Self Defense.

The classification process is aided by relatively direct relations between attributes of the decision categories and the components of the episode schema. The criminal law has evolved so that the main attributes of the decision categories suggested by legal experts (Kaplan, 1978)-- identify, mental state, circumstances, and actions -- correspond closely to the central features of human action sequences represented as episodes-- initiating events, goals, actions, and accompanying states.

The story classification stage also involves the application of the judge's procedural instructions on the presumption of innocence and the standard of proof. That is, if not all of the verdict attributes for a given verdict category are satisfied "beyond a reasonable doubt," by events in the accepted story, then the juror should presume innocence and return a default verdict of not guilty.

Confidence in Decisions. Several aspects of the decision process influence the juror's level of certainty about the final decision. First, the accepted story is judged to be the most coherent but the level of coherence will affect confidence. Thus, if the story lacks completeness, consistency, or plausibility, confidence in the story and therefore in the verdict will be diminished. Second, if a story lacks uniqueness, that is, there is more than one coherent story, then certainty concerning the accuracy of any one explanation will be lowered (Einhorn & Hogarth, 1986). Finally, the goodness-of-fit between the accepted story and the best-fitting verdict category will influence confidence in the verdict decision.

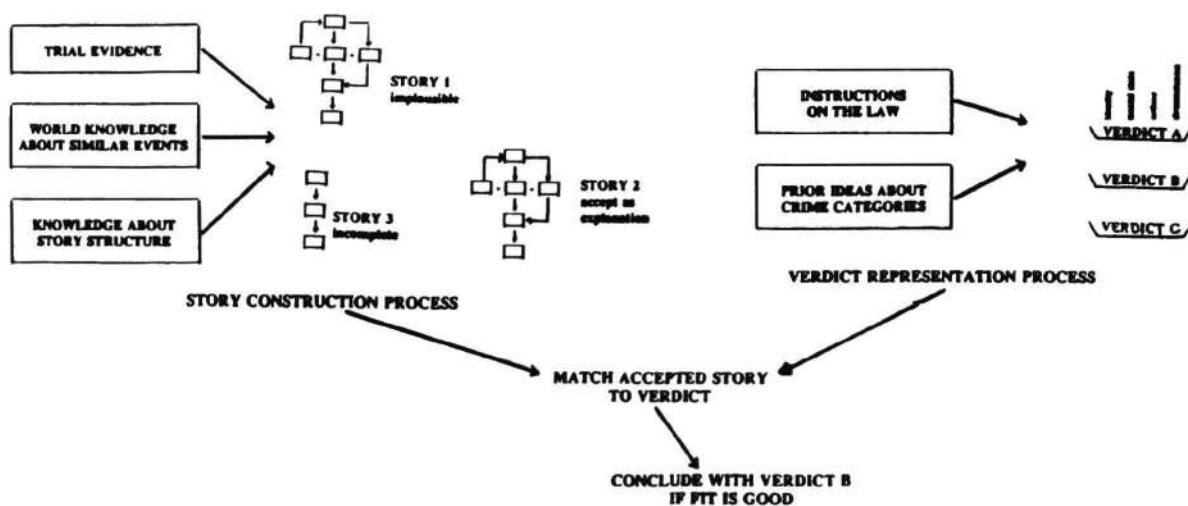


Figure 2
The "Story Model" for Juror Decision Making

In summary, our application of the general explanation-based decision model to legal decisions is based on the hypothesis that jurors impose a narrative story organization on trial information, in which causal and intentional relations between events are central (Bennett & Feldman, 1981; Pennington, 1981; Pennington & Hastie, 1986). Meaning is assigned to trial evidence through the incorporation of that evidence into one or more plausible accounts or stories describing "what happened" during events testified to at the trial. The story organization facilitates evidence comprehension and enables jurors to reach a predeliberation verdict decision. We call our application The Story Model because of the central role played by narrative, story-like evidence summaries in the decision process. The Story Model includes three components: (a) evidence evaluation through story construction, (b) representation of the decision alternatives by learning verdict category attributes, and (c) reaching a decision through the classification of the story into the best fitting verdict category (see Figure 2).

Previous Research

Our previous research on the Story Model provided descriptions of mental representations of evidentiary information and verdict information at one point in time during the decision process (Pennington & Hastie, 1986). In that research we established that the evidence summaries constructed by jurors had story structure (and not other plausible structures); verdict representations looked like feature lists (or simple frames); and that jurors who chose different verdicts had constructed different stories such that there was a distinct causal configuration of events that constituted a story corresponding to each verdict category. Moreover, jurors choosing different verdicts did not have systematically different verdict representations, nor did they apply different classification criteria. Thus verdict decisions covary with story structures but do not covary with verdict learning or story classification. However, the interview method used in this research precluded strong inferences concerning the spontaneity of story construction the functional role of stories in the decision phase.

In a second empirical study we established that decision makers spontaneously constructed causal accounts of the evidence in the legal decision task (Pennington & Hastie, 1987). In this study, subjects' responses to sentences presented in a recognition memory task were used to draw conclusions about subjects' post-decision representations of evidence. Subjects were expected to "recognize" as having been presented as trial evidence sentences from the story associated with their decision, with a higher probability than to recognize sentences from stories associated with other (rejected) decisions. This implies that hit rates (correct recognitions) and false alarm rates (false recognitions) for sentences from each story can be predicted from subjects' verdicts. These predictions were confirmed; verdict decisions predicted the high hit and false alarm rates found for sentences in the subjects' stories. Thus, a different method, subject population, and stimulus materials yielded results converging with the interview study conclusions about the correlation between memory structure and decision outcome. Even though we can conclude that story representations were constructed spontaneously, the causal role of stories in decisions is still not established because subjects could decide on a verdict and then (spontaneously) justify it to themselves by constructing a coherent story.

The Effect of Evidence Structure on Decisions

An experiment was conducted to study the effects of variations in the order of evidence presentation on judgments. Our primary goal was to test the claim that the construction of stories in evidence evaluation causes decisions. A secondary goal was to determine whether story coherence and uniqueness influence judgments of confidence in the correctness of verdicts.

We used legal case materials based on the transcript of an actual murder trial, titled Commonwealth v. Johnson, and varied presentation order to influence the ease with which a prosecution (First Degree Murder) or defense (Not Guilty by Reason of Self Defense) story could be constructed (see Pennington & Hastie, 1986 for a summary of the trial). The "logic" of the experiment was summarized in our hypothesis that (manipulated) ease of story construction would influence verdict decisions; easy-to-construct stories would result in more decisions in favor of the corresponding verdicts.

Stories were considered easy to construct when the evidence was ordered in a temporal and causal sequence that matched the occurrence of the original events (Story Order; Baker, 1978). Stories were considered difficult to construct when the presentation order did not match the sequence of the original events. We based the non-story order on the sequence of evidence as conveyed by witnesses in the original trial (Witness Order). One-hundred and thirty college student mock-jurors listened to a tape recording of a 100-sentence summary of the trial evidence (50 prosecution statements and 50 defense statements), followed by a judge's charge to choose between a Murder verdict and a Not Guilty verdict. The 50 prosecution statements, constituting the First Degree Murder story identified in our initial interview study (Pennington & Hastie, 1986), were presented either in a Story Order or a Witness Order. Similarly, the defense statements, the Not Guilty story, were presented in one of the two orders creating a four-cell factorial design. In all four order conditions the prosecution evidence preceded the defense evidence as per standard legal procedure. After listening to the tape recorded trial materials, the subjects completed a questionnaire indicating their verdict, confidence in the verdict, and their perceptions of the strengths of the prosecution and defense cases.

As predicted, subjects were likeliest to convict the defendant when the prosecution evidence was presented in Story Order and the defense evidence was presented in Witness Order (78% chose guilty) and they were least likely to convict when the prosecution evidence was in Witness Order and defense was in Story Order (31% chose guilty, see Table 1). Conviction rates were intermediate in conditions where both sides of the case were in Story Order (59% convictions) or both were in Witness Order (63% convictions). Statistically, the best summary of the effects of evidence order on verdict choice was two main effects, one for defense side order (Story versus Witness) and one for prosecution side order (log-linear model analysis, chi-squared "badness-of-fit" statistic (2 df) = .42, $p > .80$).

Analyses were conducted on the ratings of strength of the defense and prosecution cases and these ratings were influenced by presentation order, with Story Order evidence rated as stronger than Witness Order. Furthermore, the perceived strength of one side of the case depended on both the order of evidence for that side and for the other side of the case. This finding supports our claim that the uniqueness of the best-fitting story is one important basis for confidence in the decision. We also examined the verdict confidence ratings and found that, regardless of verdict chosen, jurors who heard both sides of the case in Story Order were

verdict chosen, jurors who heard both sides of the case in Story Order were more confident than jurors who heard one or neither side in Story Order. This result reinforces our conclusion that alternate story strength is important, although the finding was not predicted.

		Defense Case Presented In	
		<u>STORY FORM</u>	<u>WITNESS FORM</u>
Prosecution Case Presented In	STORY FORM	59%	78%
	WITNESS FORM	31%	63%

Table 1
Percentages of Subjects Choosing the "Guilty" Verdict

Conclusions

We have introduced a model of decision making that describes human behavior in tasks where a large, implication-rich, conditionally-dependent set of propositions constitute an evidence base for selection of an option from a limited set of decision alternatives. We propose that considerable processing occurs to understand the evidence base by constructing a summary explanation or causal model of the decision-relevant facts, assumptions, and premises. Once a satisfactory explanation has been constructed, the decision maker attempts to select an option by matching features of the summary explanation to corresponding characteristics of solutions or courses of action in the decision set.

The decision process is divided into three stages: construction of a summary explanation; determination of decision alternatives; mapping the explanation onto a best-fitting decision alternative. This subtask framework is in contrast to the uniform on-line updating computation or the unitary memory-based calculation hypothesized in most alternative approaches (cf. Hastie & Park, 1986). Furthermore, we diverge sharply from traditional approaches with our emphasis on the structure of memory representations as the key determinant of decisions. We also depart from the common assumption that, when causal reasoning is involved in judgment, it can be described by algebraic, stochastic, or logical computations that lead directly to a decision (e.g., Anderson, 1974; Einhorn & Hogarth, 1985; Kelley, 1973). In our model causal reasoning plays a subordinate but critical role by guiding inferences in evidence evaluation and construction of the intermediate explanation.

References

- Anderson, N.H. (1974). Cognitive algebra: Integration theory applied to social attribution. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology, Volume 7. New York: Academic.
- Anderson, N.H. (1981). Foundations of Information Integration Theory. New York: Academic.
- Axelrod, R. (Ed.) (1976). Structure of Decision: The Cognitive Maps of Political Elites. Princeton: Princeton University Press.
- Baker, L. (1978). Processing temporal relationships in simple stories: Effects of input sequences. Journal of Verbal Learning and Verbal Behavior, 17, 559-572.
- Bennett, W.L., & Feldman, M. (1981). Reconstructing Reality in the Courtroom. New Brunswick, N.J.: Rutgers University Press.
- deKleer, J., & Brown, J.S. (1983). Assumptions and ambiguities in mechanistic mental models. In D. Gentner and A.L. Stevens (Eds.), Mental Models. Hillsdale, N.J.: Erlbaum.
- Edwards, W. (1954). The theory of decision making. Psychological Review, 51, 380-417.
- Einhorn, H.J., & Hogarth, R.M. (1986). Judging probable cause. Psychological Bulletin, 99, 3-19.
- Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. Psychological Review, 93, 258-268.
- Hogarth, R.M., Michaud, C., & Mery, J.L. (1980). Decision behavior in urban development: A methodological approach and substantive considerations. Acta Psychologica, 45, 95-117.
- Hutchins, E. (1980). Culture and Inference. Cambridge, Mass.: Harvard University Press.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47, 263-291.
- Kaplan, J. (1978). Criminal Justice: Introductory Cases and Materials (2nd Ed.). Mineola, N.Y.: Foundation Press.
- Kelley, H.H. (1973). The processes of causal attribution. American Psychologist, 28, 107-128.
- Mandler, J.M. (1980). Categorical and schematic organization in memory. In C.R. Puff (Ed.), Memory Organization and Structure. New York: Academic.
- Pennington, N. (1981). Causal Reasoning and Decision Making: The Case of Juror Decisions. Unpublished doctoral dissertation, Harvard University.

Pennington, N., & Hastie, R. (1981). Juror decision making models: The generalization gap. Psychological Bulletin, 89, 246-287.

Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. Journal of Personality and Social Psychology, 51, 242-258.

Pennington, N., & Hastie, R. (1987). Explanation-based decision making: The effects of memory structure on judgment. Unpublished manuscript, University of Chicago.

Pople, H.E., Jr. (1982). Heuristic methods for imposing structure on ill-structured problems: The structuring of medical diagnostics. In P. Szolovits (Ed.), Artificial Intelligence in Medicine. Boulder, Colo.: Westview Press.

Rumelhart, D.E. (1977). Understanding and summarizing brief stories. In D. LaBerge and S.J. Samuels (Eds.), Basic Processes in Reading: Perception and Comprehension. Hillsdale, N.J.: Erlbaum.

Stein, N.L., & Glenn, C.G. (1979). An analysis of story comprehension in elementary school children. In R.O. Freedle (Ed.), New Directions in Discourse Processing, Volume 2. Norwood, N.J.: Ablex.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. Journal of Memory and Language, 24, 612-630.