Cognition, 50 (1994) 403–430

Uncertainty and the difficulty of thinking through disjunctions

Eldar Shafir*
Department of Psychology, Princeton University, Princeton, NJ 08544, USA

Abstract

This paper considers the relationship between decision under uncertainty and thinking through disjunctions. Decision situations that lead to violations of Savage's sure-thing principle are examined, and a variety of simple reasoning problems that often generate confusion and error are reviewed. The common difficulty is attributed to people's reluctance to think through disjunctions. Instead of hypothetically traveling through the branches of a decision tree, it is suggested, people suspend judgement and remain at the node. This interpretation is applied to instances of decision making, information search, deductive and inductive reasoning, probabilistic judgement, games, puzzles and paradoxes. Some implications of the reluctance to think through disjunctions, as well as potential corrective procedures, are discussed.

Introduction

Everyday thinking and decision making often occur in situations of uncertainty. A critical feature of thinking and deciding under uncertainty is the need to consider possible states of the world and their potential consequences for our beliefs and actions. Uncertain situations may be thought of as disjunctions of possible states: either one state will obtain, or another. In order to choose between alternative actions or solutions in situations of uncertainty, a person

*E-mail eldar@clarity.princeton.edu
This research was supported by US Public Health Service Grant No. 1-R29-MH46885 from the National Institute of Mental Health, and by a grant from the Russell Sage Foundation. The paper has benefited from long discussions with Amos Tversky, and from the comments of Philip Johnson-Laird, Daniel Osherson, and Edward Smith.

SSDI 0010-0277(93)00612-B

NIMH

needs to consider the anticipated outcomes of each action or each solution pattern under each state. Thus, when planning a weekend's outing, a person may want to consider which of a number of activities she would prefer if the weekend is sunny and which she would prefer if it rains. Similarly, when contemplating the next move in a chess game, a player needs to consider what the best move would be if the opponent were to employ one strategy, and what may be the best move if the opponent were to follow an alternative plan.

Special situations sometimes arise in which a particular action, or solution, yields a more desirable outcome no matter how the uncertainty is resolved. Thus, a person may prefer to go bowling rather than hiking regardless of whether it is sunny or it rains, and an exchange of queens may be the preferred move whatever the strategy chosen by the opponent. An analogous situation was described by Savage (1954) in the following passage:

> A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant to the attractiveness of the purchase. So, to clarify the matter for himself, he asks whether he would buy if he knew that the Republican candidate were going to win, and decides that he would do so. Similarly, he considers whether he would buy if he knew that the Democratic candidate were going to win, and again finds that he would do so. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains . . .

Savage calls the principle that governs this decision the sure-thing principle (STP). According to STP, if a person would prefer a to b knowing that X obtained, and if he would also prefer a to b knowing that X did not obtain, then he definitely prefers a to b (Savage, 1954, p. 22). STP has a great deal of both normative and descriptive appeal, and is one of the simplest and least controversial principles of rational behavior. It is an important duplication of "consequentialist" accounts of decision making, in that it captures a fundamental intuition about what it means for a decision to be determined by the anticipated consequences.[1] It is a cornerstone of expected utility theory, and it holds in other models of choice which impose less stringent criteria of rationality (although see McClennen, 1983, for discussion). Despite its apparent simplicity, however, people's decisions do not always abide by STP.

The present paper reviews recent experimental studies of decision under uncertainty that exhibit violations of STP in simple disjunctive situations. It is argued that a necessary condition for such violations is people's failure to see through the underlying disjunctions. In particular, it is suggested that in situations of uncertainty people tend to refrain from fully contemplating the consequences of potential outcomes and, instead, suspend judgement and remain, undecided, at

[1] The notion of consequentialism appears in the decision theoretic literature in a number of different senses. See, for example, Hammond (1988), Levi (1991), and Bacharach and Hurley (1991) for technical discussion. See also Shafir and Tversky (1992) for a discussion of nonconsequentialism.

the uncertain node. Studies in other areas, ranging from deduction and probability judgement to games and inductive inference, are then considered, and it is argued that a reluctance to think through disjunctions can be witnessed across these diverse domains. Part of the difficulty in thinking under uncertainty, it is suggested, derives form the fact that uncertainty requires thinking through disjunctive situations. Some implications and corrective procedures are considered in a concluding section.

Decisions

Risky choice

Imagine that you have just gambled on a toss of a coin in which you had an equal chance to win $200 or lose $100. Suppose that the coin has been tossed, but that you do not know whether you have won or lost. Would you like to gamble again, on a similar toss? Alternatively, how would you feel about taking the second gamble given that you have just lost $100 on the first (henceforth, the *Lost* version)? And finally, would you play again given that you have won $200 on the first toss (the *Won* version)?

Tversky and Shafir (1992) presented subjects with the Won, Lost, and uncertain versions of this problem, each roughly a week apart. The problems were embedded among several others so the relation among the three versions would not be transparent, and subjects were instructed to treat each decision separately. The data were as follows: the majority of subjects accepted the second gamble after having won the first gamble, the majority accepted the second gamble after having lost the first gamble, but most subjects rejected the second gamble when the outcome of the first was not known. Among those subjects who accepted the second gamble both after a gain and after a loss on the first, 63% rejected the second gamble in the disjunctive condition, when the outcome of the first gamble was uncertain. In fact, this particular pattern – accept when you win, accept when you lose, but reject when you do not know – was the single most frequent pattern exhibited by our subjects (see Tversky & Shafir, 1992, for further detail and related data).

A decision maker who would choose to accept the second gamble both after having won and after having lost the first, should – in conformity with STP – choose to accept the second gamble even when the outcome of the first is uncertain. However, when it is not known whether they have won or lost, our subjects refrain from contemplating (and acting in accordance with) the consequences of winning or of losing. Instead, they act as if in need for the uncertainty about the first toss to be resolved. Elsewhere, we have suggested that people have

different reasons for accepting the second gamble following a gain and following a loss, and that a disjunction of different reasons (" 'I can no longer lose . . .' in case I won the first gamble or 'I need to recover my losses . . .' in case I lost") is often less compelling than either of these definite reasons alone (for further discussion of the role of reasons in choice, see Shafir, Simonson, & Tversky, 1993).

Tversky and Shafir (1992) call the above pattern of decisions a *disjunction effect*. A disjunction effect occurs when a person prefers x over y when she knows that event A obtains, and she also prefers x over y when she knows that event A does not obtain, but she prefers y over x when it is unknown whether or not A obtains. The disjunction effect amounts to a violation of STP, and hence of consequentialism. While a reliance on reasons seems to play a significant role in the psychology that yields disjunction effects, there is nonetheless another important element that contributes to these paradoxical results: people do not see through the otherwise compelling logic that characterizes these situations. When confronting such disjunctive scenarios, which can be thought of as decision trees, people seem to remain at the uncertain nodes, rather then contemplate the – sometimes incontrovertible – consequences of the possible branches.

The above pattern of nonconsequential reasoning may be illustrated with the aid of the value function from Kahneman and Tversky's (1979) prospect theory. The function, shown in Fig. 1, represents people's subjective value of losses and of gains, and captures common features of preference observed in numerous empirical studies. Its S-shape combines a concave segment to the right of the origin reflecting risk aversion in choices between gains, and a convex segment to the left of the origin reflecting risk seeking in choices between losses. Furthermore, the slope of the function is steeper on the left of the origin than on the right, reflecting the common observation that "losses loom larger than gains" for most people. (For more on prospect theory, see Kahneman & Tversky, 1979, 1982, as well as Tversky & Kahneman, 1992, for recent extensions.) The function in Fig. 1 represents a typical decision maker who is indifferent between a 50% chance of winning \$100 and a sure gain of roughly \$35, and, similarly, is indifferent between a 50% chance of losing \$100 and a sure loss of roughly \$40. Such a pattern of preferences can be captured by a power function with an exponent of .65 for gains and .75 for losses. While prospect theory also incorporates a decision weight function, π, which maps stated probabilities into their subjective value for the decision maker, we will assume, for simplicity, that decision weights coincide with stated probabilities. While there is ample evidence to the contrary, this does not change the present analysis.

Consider, then, a person P whose values for gains and losses are captured by the function of Fig. 1. Suppose that P is presented with the gamble problem above and is told that he has won the first toss. He now needs to decide whether to accept or reject the second. P needs to decide, in other words, whether to

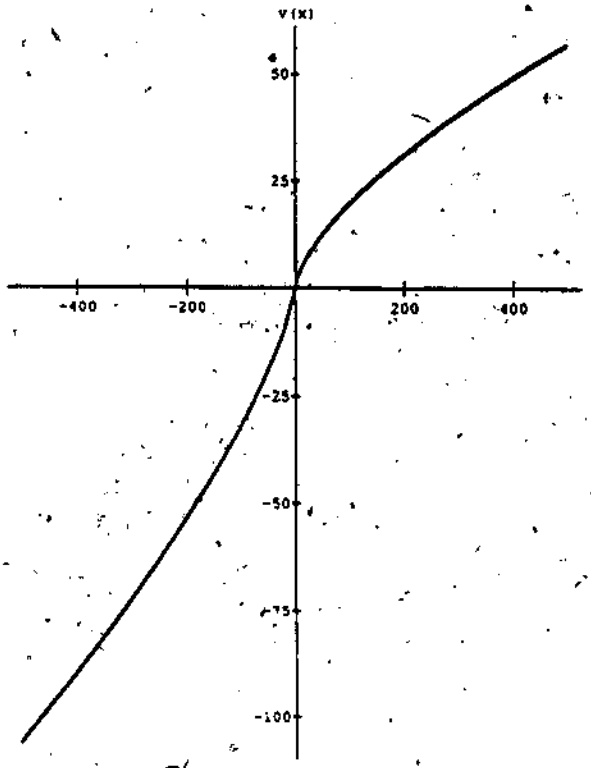


Fig. 1. *The value function* $v(x) = x^{(.65)}$ *for* $x \geq 0$ *and* $v(x) = -(-x)^{(.65)}$ *for* $x \leq 0$.

maintain a sure gain of $200 or, instead, opt for an equal chance at either a $100 or a $400 gain. Given P's value function, his choice is between two options whose expected values are as follows:

Accept the second gamble: $.50 \times 400^{(.65)} + .50 \times 100^{(.65)} = 250$
Reject the second gamble: $1.0 \times 200^{(.65)} = 200$

Because the value of the first option is greater than that of the second, P is predicted to accept the second gamble. Similarly, when P is told that he has lost the first gamble and needs to decide whether to accept or reject the second, P faces the following options:

losing twice

Accept the second gamble: $.50 \times -[200^{(.75)}] + .50 \times 100^{(.65)} = -50$
Reject the second gamble: $1.0 \times -[100^{(.75)}] = -100$

coolest ? Can I afford to loose more I want to buy X that costs etc. $200.00

Again, since the first quantity is larger than the second, P is predicted to accept the second gamble.

Thus, once the outcome of the first gamble is known, the value function of Fig. 1 predicts that person P will accept the second gamble whether he has won or lost the first. But what is P expected to do when the outcome of the first gamble is not known? Because he does not know the outcome of the first gamble, P may momentarily assume that he is still where he began – that, for moment, no changes have transpired. Not knowing whether he has won or lost, P remains for now at the status quo, at the origin of his value function. When presented with the decision to accept or reject the second gamble, P evaluates it from his original position, without incorporating the outcome of the first gamble, which remains unknown. Thus, P needs to choose between accepting or rejecting a gamble that offers an equal chance to win \$200 or lose \$100:

Accept the second gamble: $.50 \times -[100^{(.75)}] + .50 \times 200^{(.65)}$

Reject the second gamble: 0.

Because the expected value of accepting is just below 0, P decides to reject the second gamble in this case.

Thus, aided by prospect theory's value function, we see how a decision maker's "suspension of judgement" – his tendency to assume himself at the origin, or status quo, when it is not known whether he has won or lost – leads him to reject an option that he would accept no matter what his actual position may be. Situated at a chance node whose outcome is not known, P's reluctance to consider each of the hypothetical branches leads him to behave in a fashion that conflicts with his preferred behavior given either branch. People in these situations seem to confound their epistemic uncertainty – what they may or may not know – with uncertainty about the actual consequences – what may or may not have occurred. A greater focus on the consequences would have helped our subjects realize the implications for their preference of either of the outcomes. Instead, not knowing which was the actual outcome, our subjects chose to evaluate the situation as if neither outcome had obtained. It is this reluctance to think through disjunctions that characterizes many of the phenomena considered below.

Search for noninstrumental information: the Hawaiian vacation

Imagine that you have just taken a tough qualifying exam. It is the end of the semester, you feel tired and run-down, and you are not sure that you passed the exam. In case you failed you have to take it again in a couple of months – after

the Christmas holidays. You now have an opportunity to buy a very attractive 5-day Christmas vacation package to Hawaii at an exceptionally low price. The special offer expires tomorrow, while the exam grade will not be available until the following day. Do you buy the vacation package?

This question was presented by Tversky and Shafir (1992) to Stanford University undergraduate students. Notice that the outcome of the exam will be known long before the vacation begins. Thus, the uncertainty characterizes the present, disjunctive situation, not the eventual vacation. Additional, related versions were presented in which subjects were to assume that they had passed the exam, or that they had failed, before they had to decide about the vacation. We discovered that many subjects who would have bought the vacation to Hawaii if they were to pass the exam *and* if they were to fail, chose not to buy the vacation when the exam's outcome was not known. The data show that more than half of the students chose the vacation package when they knew that they passed the exam and an even larger percentage chose the vacation when they knew that they failed. However, when they did not know whether they had passed or failed, less than one-third of the students chose the vacation and the majority (61%) were willing to pay $5 to postpone the decision until the following day, when the results of the exam would be known.[1] Note the similarity of this pattern to the foregoing gamble scenario: situated at a node whose outcome is uncertain, our students envision themselves at the status quo, as if no exam had been taken. This "suspension of judgement" – the reluctance to consider the possible branches (having either passed or failed the exam) – leads our subjects to behave in a manner that conflicts with their preferred option given either branch.

The pattern observed in the context of this decision is partly attributed by Tversky and Shafir (1992) to the reasons that subjects summon for buying the vacation (see also Shafir, Simonson, & Tversky, 1993, for further discussion). Once the outcome of the exam is known, the student has good – albeit different – reasons for going to Hawaii: having passed the exam, the vacation can be seen as a reward following a successful semester; having failed the exam, the vacation becomes a consolation and time to recuperate before a re-examination. Not knowing the outcome of the exam, however, the student lacks a definite reason for going to Hawaii. The indeterminacy of reasons discourages many students from buying the vacation, even when both outcomes – passing or failing the exam – ultimately favor this course of action. Evidently, a disjunction of different

[1] Another group of subjects were presented with both Fail and Pass versions, and asked whether they would buy the vacation package in each case. Two-thirds of the subjects made the same choice in the two conditions, indicating that the data for the disjunctive version cannot be explained by the hypothesis that those who buy the vacation in case they pass the exam do not buy it in case they fail, and vice versa. While only one-third of the subjects made different decisions depending on the outcome of the exam, more than 60% of the subjects chose to wait when the outcome was not known.

reasons (reward in case of success; consolation in case of failure) can be less compelling than either definite reason alone.

A significant proportion of subjects were willing to pay, in effect, for information that was ultimately not going to alter their decision – they would choose to go to Hawaii in either case. Such willingness to pay for noninstrumental information is 'at variance' with the classical model, in which the worth of information is determined by its potential to influence choice. People's reluctance to think through disjunctive situations, on the other hand, entails that noninstrumental information will sometimes be sought. (See Bastardi & Shafir, 1994, for additional studies of the search for noninstrumental information and its effects on choice.) While vaguely aware of the possible outcomes, people seem reluctant to fully entertain the consequences as long as the actual outcome is uncertain. When seemingly relevant information may become available, they often prefer to have the uncertainty resolved, rather than consider the consequences of each branch of the tree under the veil of uncertainty. A greater tendency to consider the potential consequences may sometimes help unveil the noninstrumental nature of missing information. In fact, when subjects were first asked to contemplate what they would do in case they failed the exam and in case they passed, almost no subject who had expressed the same preference for both outcomes then chose to wait to find out which outcome obtained (Tversky & Shafir, 1992).

The decision of many subjects in the disjunctive scenario above was not guided by a simple evaluation of the consequences (for, then, they would have realized that they prefer to go to Hawaii in either case). An adequate account of this behavior needs to contend with the fact that the very simple and compelling disjunctive logic of STP does not play a decisive role in subjects' reasoning. A behavioral pattern which systematically violates a simple normative rule requires both a positive as well as a negative account (see Kahneman and Tversky, 1982, for discussion). We need to understand not only the factors that produce a particular response, but also why the correct response is not made. Work on the conjunction fallacy (Shafir, Smith, & Osherson, 1990; Tversky and Kahneman, 1983), for example, has addressed both the fact that people's probability judgement relies on the representativeness heuristic – a positive account – as well as the fact that people do not perceive the extensional logic of the conjunction rule as decisive – a negative account. The present work focuses on the negative facet of nonconsequential reasoning and STP violations. It argues that like other principles of reasoning and decision making, STP is very compelling when stated in a general and abstract form, but is often non-transparent, particularly because it applies to disjunctive situations. The following section briefly reviews studies of nonconsequential decision making in the context of games, and ensuing sections extend the analysis to other domains.

Games

Prisoner's dilemma

The theory of games explores the interaction between players acting according to specific rules. One kind of two-person game that has received much attention is the Prisoner's dilemma, or PD. (For an extensive treatment, see Rapoport & Chammah, 1965). A typical PD is presented in Fig. 2.

The cell entries indicate the number of points each player receives contingent on the two players' choices. Thus, if both cooperate each receives 75 points but if, for example, the other cooperates and you compete, you receive 85 points while the other receives 25. What characterizes the PD is that no matter what the other does, each player fares better if he competes than if he cooperates; yet, if they both compete they do significantly less well than if they had both cooperated. Since each player is encountered at most once, there is no opportunity for conveying strategic messages, inducing reciprocity, or otherwise influencing the other player's choice of strategy.

A player in a PD faces a disjunctive situation. The other chooses one of two strategies, either to compete or to cooperate. Not knowing the other's choice, the first player must decide on his own strategy. Whereas each player does better competing, their mutually preferred outcome results from mutual cooperation rather than competition. A player, therefore, experiences conflicting motivations. Regardless of what the other does, he is better off being selfish and competing; but assuming that the other acts very much like himself, they are better off both making the ethical decision to cooperate rather than the selfish choice to compete. How might this disjunctive situation influence people's choice of strategy?

		OTHER: cooperates	OTHER: competes
YOU	cooperate	You: 75 Other: 75	You: 25 Other: 85
	compete	You: 85 Other: 25	You: 30 Other: 30

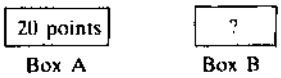
Fig. 2. *A typical prisoner's dilemma. The cell entries indicate the number of points that you and the other player receive contingent on your choices.*

Shafir and Tversky (1992) have documented disjunction effects in one-shot PD games played for real payoffs. Subjects ($N = 80$) played a series of PD games (as in Fig. 2) on a computer, each against a different unknown opponent supposedly selected at random from among the participants. Subjects were told that they had been randomly assigned to a "bonus group", and that occasionally they would be given information about the other player's already-chosen strategy before they had to choose their own. This information appeared on the screen next to the game, and subjects were free to take it into account in making their decision. (For details and the full instructions given to subjects, see Shafir & Tversky, 1992.) The rate of cooperation in this setting was 3% when subjects knew that the opponent had defected, and 16% when they knew that the opponent had cooperated. Now what should subjects do when the opponent's decision is not known? Since 3% cooperate when the other competes and 16% cooperate when the other cooperates, one would expect an intermediate rate of cooperation when the other's strategy is not known. Instead, when subjects did not know whether their opponent had cooperated or defected (as is normally the case in this game), the rate of cooperation rose to 37%. In violation of STP, a quarter of the subjects defected when they knew their opponent's choice – be it cooperation or defection – but cooperated when their opponent's choice was not known. Note the recurring pattern: situated at a disjunctive node whose outcome is uncertain, these subjects envision themselves at the status quo; as if, for the moment, the uncertain strategy selected by the opponent has no clear consequences. These players seem to confound their epistemic uncertainty – what they may or may not know about the other's choice of strategy – with uncertainty about the actual consequences – the fact that the other is bound to be a cooperator or a defector, and that they, in turn, are bound to respond by defecting in either case. (For further analysis and a positive account of what may be driving subjects' tendency to cooperate under uncertainty, see Shafir & Tversky, 1992.)

Newcomb's problem and quasi-magical thinking

Upon completing the PD game described in the previous section, subjects ($N = 40$) were presented, on a computer screen, with the following scenario based on the celebrated Newcomb's problem (for more on Newcomb's problem, see Nozick, 1969; see Shafir & Tversky, 1992, for further detail and discussion of the experiment).

> You now have one more chance to collect additional points. A program developed recently at MIT was applied during this entire session to analyze the pattern of your preferences. Based on that analysis, the program has predicted your preference in this final problem.



Consider the two boxes above. Box A contains 20 points for sure. Box B may or may not contain 250 points. Your options are to:

(1) Choose both boxes (and collect the points that are in both).
(2) Choose Box B only (and collect only the points that are in Box B).

If the program predicted, based on observation of your previous preferences, that you will take both boxes, then it left Box B empty. On the other hand, if it predicted that you will take only Box B, then it put 250 points in that box. (So far, the program has been remarkably successful: 92% of the participants who choose only Box B found 250 points in it, as opposed to 17% of those who chose both boxes.)

To insure that the program does not alter its guess after you have indicated your preference, please indicate to the person in charge whether you prefer both boxes or Box B only. After you indicate your preference, press any key to discover the allocation of points.

According to one rationale that arises in the context of this decision, if the person chooses both boxes, then the program, which is remarkably good at predicting preferences, is likely to have predicted this and will not have put the 250 points in the opaque box. Thus, the person will get only 20 points. If, on the other hand, the person takes only the opaque box, the program is likely to have predicted this and will have put the 250 points in that box, and so the person will get 250 points. A subject may thus be tempted to reason that if he takes both boxes he is likely to get only 20 points, but that if he takes just the opaque box he is likely to get 250 points. There is a compelling motivation to choose just the opaque box, and thereby resemble those who typically find 250 points it it.

There is, of course, another rationale: the program has already made its prediction and has already either put the 250 points in the opaque box or has not. If it has already put the 250 points in the opaque box, and the person takes both boxes he gets 250 + 20 points, whereas if he takes only the opaque box, he gets only 20 points. If the program has not put the 250 points in the opaque box and the person takes both boxes he gets 20 points, whereas if he takes only the opaque box he gets nothing. Therefore, whether the 250 points are there or not, the person gets 20 points more by taking both boxes rather than the opaque box only.

The second rationale relies on consequentialist reasoning reminiscent of STP (namely, whatever the state of the boxes following the program's prediction, I will do better choosing both boxes rather than one only). The first rationale, on the other hand, while couched in terms of expected value, is partially based on the assumption that what the program will have predicted – although it has predicted this already – depends somehow on what the subject ultimately decides to do.

The results we obtained were as follows: 35% of the subjects chose both boxes, while 65% preferred to take Box B only. This proportion of choices is similar to that observed in other surveys concerning the original Newcomb's problem (see,

for example, Gardner, 1973, 1974; Hofstadter, 1983). What can be said about the majority who prefer to take just one box? Clearly, had they known for certain that there were 250 points in the opaque box (and could see 20 in the other), they would have taken both rather than just one. And certainly, if they knew that the 250 points were not in that box, they would have taken both rather than just the one that's empty. These subjects, in other words, would have taken both boxes had they known that Box B is either full or empty, but a majority preferred to take only Box B when its contents were not known.

The conflicting intuitions that subjects experience in the disjunctive situation – when the program's prediction is not known – are obviously resolved in favor of both boxes once the program's decision has been announced: at that point, no matter what the program has predicted, taking both boxes brings more points. Subjects, therefore, should choose both boxes also when the program's decision is uncertain. Instead, many subjects fail to be moved by the foreseeable consequences of the program's predictions, and succumb to the strong motivation to choose just the opaque box and thereby resemble those who typically find 250 points in it.[3] As Gibbard and Harper (1978) suggest in an attempt to explain people's choice of a single box, "a person may . . . want to bring about an indication of a desired state of the world, even if it is known that the act that brings about the indication in no way brings about the desired state itself". This form of magical thinking was demonstrated by Quattrone and Tversky (1984), whose subjects selected actions that were diagnostic of favorable outcomes even though the actions could not cause those outcomes. Note that such instances of magical thinking typically occur in disjunctive situations, before the exact outcome is known. Once they are aware of the outcome, few people think they can reverse it by choosing an action that is diagnostic of an alternative event.

Shafir and Tversky (1992) discuss various manifestations of "quasi-magical" thinking, related to phenomena of self-deception and illusory control. These include people's tendency to place larger bets before rather than after a coin has been tossed (Rothbart & Snyder, 1970; Strickland, Lewicki, & Katz, 1966), or to throw dice softly for low numbers and harder for high ones (Henslin, 1967). Similarly, Quattrone and Tversky (1984) note that Calvinists act as if their behavior will determine whether they will go to heaven or to hell, despite their belief in divine pre-determination, which entails that their fate has been determined at birth. The presence of uncertainty, it appears, is a major contributor to quasi-magical thinking; few people act as if they can undo an already certain

[3]The fact that subjects do not see through this disjunctive scenario seems indisputable. It is less clear, however, what conditions would serve to make the situation more transparent, and to what extent. Imagine, for example, that subjects were given a sealed copy of the program's decision to take home with them, and asked to inspect it that evening, after having made their choice. It seems likely that an emphasis on the fact that the program's decision has long been made would reduce the tendency to choose a single box.

event, but while facing a disjunction of events, people often behave as if they can exert some control over the outcome. Thus, many people who are eager to vote while the outcome is pending, may no longer wish to do so once the outcome of the elections has been determined. In this vein, it is possible that Calvinists would perhaps do fewer good deeds if they knew that they had already been assigned to heaven, or to hell, than while their fate remains a mystery. Along similar lines, Jahoda (1969) discusses the close relationship between uncertainty and superstitious behavior, which is typically exhibited in the context of uncertain outcomes rather than in an attempt to alter events whose outcome is already known.

As illustrated by the studies above, people often are reluctant to consider the possible outcomes of disjunctive situations, and instead suspend judgement and envision themselves at the uncertain node. Interestingly, it appears that decision under uncertainty is only one of numerous domains in which subjects exhibit a reluctance to think through disjunctive situations. The difficulties inherent to thinking through uncertainty and, in particular, people's reluctance to think through disjunctions manifest themselves in other reasoning and problem-solving domains, some of which are considered below.

Probabilistic judgement

Researchers into human intuitive judgement as well as teachers of statistics have commented on people's difficulties in judging the probabilities of disjunctive events (see, for example, Bar-Hillel, 1973; Carlson & Yates, 1989; Tversky & Kahneman, 1974). While some disjunctive predictions may in fact be quite complicated, others are simple, assuming that one sees though their disjunctive character. Consider, for example, the following "guessing game" which consisted of two black boxes presented to Princeton undergraduates ($N = 40$) on a computer screen, along with the following instructions.

> Under the black cover, each of the boxes above is equally likely to be either white, blue, or purple. You are now offered to play one of the following two games of chance:
>
> *Game 1:* You guess the color of the left-hand box. You win 50 points if you were right, and nothing if you were wrong.
>
> *Game 2:* You choose to uncover both boxes. You win 50 points if they are the same color, and nothing if they are different colors.

The chances of winning in Game 1 are 1/3; the chances of winning in Game 2 are also 1/3. To see that, one need only realize that the first box is bound to be either white, blue, or purple and that, in either case, the chances that the other will be the same color are 1/3. Notice that this reasoning incorporates the disjunctive

logic of STP. One enumerates the possible outcomes of the first box, considers the chances of winning conditional on each outcome, and realizes that the chances are the same no matter what the first outcome was. Subjects, therefore, are expected to find the two games roughly equally attractive, provided that they see through the disjunctive nature of Game 2. This disjunctive rationale, however, seems not to have been entirely transparent to our subjects, 70% of whom indicated a preference for Game 1 (significantly different from chance, $Z = 2.53$, $p < .05$). These subjects may have suspected an equal chance for both games, but a certain lack of clarity about the disjunctive case may have led them to prefer the unambiguous first game. Whereas this preference could also be attributed to subjects' beliefs about the computer set-up, the next version not only insured the perceived independence of outcomes, but also emphasized the game's sequential character which, it was thought, may make its disjunctive nature more transparent. One hundred and three Stanford undergraduates listed their highest buying prices for the gambles below:

> The following games of chance are played with a regular die that has two yellow sides, two green sides, and two red sides:
>
> *Game A*: You roll the die once. You win $40 if it falls on green, and nothing otherwise. What is the largest amount of money that you would be willing to pay to participate in this game?
>
> *Game B*: You roll the die twice. You win $40 if it falls on the same color both times (e.g., both red, or both green) and nothing otherwise. What is the largest amount of money that you would be willing to pay to participate in this game?

The probability of winning in Game A is 1/3. The probability of winning in Game B is also 1/3. To see this, one need only realize that for every outcome of the first toss, the probability of winning on the second toss is always 1/3. Forty-six percent of our subjects, however, did not list the same buying price for the two games. Eighty-five percent of these subjects offered a higher price for Game A than for Game B. In fact, over all subjects, Game A was valued at an average of $6.09, while Game B was worth an average of only $4.69 ($t = 4.65$, $p < .001$). We have investigated numerous scenarios of this kind, in all of which a large proportion of subjects prefer to gamble on a simple event over an equally likely or more likely disjunctive event.

Inductive inference

Inferential situations often involve uncertainty not only about the conclusion, but about the premises as well. The judged guilt of a defendant depends on the veracity of the witnesses; the diagnosis of a patient depends on the reliability of the tests; and the truth of a scientific hypothesis depends on the precision of

earlier observations. Reasoning from uncertain premises can be thought of as reasoning through disjunctions. How likely is the defendant to be guilty if the witness is telling the truth and how likely if the witness is lying? What is the likelihood that the patient has the disease given that the test results are right, and what is the likelihood if the results are false? The aggregation of uncertainty is the topic of various theoretical proposals (see, for example, Shafer & Pearl, 1990), all of which agree on a general principle implied by the probability calculus. According to this principle, if I believe that event A is more probable than event B in light of some condition c, and if I also believe that event A is more probable than event B given the absence of c, then I believe that A is more probable than B regardless of whether c obtains or not. Similar to the failure of STP in the context of choice, however, this principle may not always describe people's actual judgements.

In the following pilot study, 182 Stanford undergraduates were presented with the divorce scenario below, along with one of the three questions that follow:

Divorce problem
Tim and Julia, both school teachers, have been married for 12 years. They have a 10-year old son, Eric, to whom they are very attached. During the last few years Tim and Julia have had recurring marital problems. They have consulted marriage counselors and have separated once for a couple of months, but decided to try again. Their marriage is presently at a new low.

Disjunctive question ($N = 88$):
What do you think are the chances that both Tim and Julia will agree to a divorce settlement (that specifies whether Eric is to stay with his father or with his mother)?
[59.8%]

Mother question ($N = 46$):
What do you think are the chances that both Tim and Julia will agree to a divorce settlement if Eric is to stay with his mother?
[49.8%]

Father question ($N = 48$):
What do you think are the chances that both Tim and Julia will agree to a divorce settlement if Eric is to stay with his father?
[40.7%]

Next to each question is its mean probability rating. Subjects judged the probability that the parents will agree to a divorce settlement that specified that the child is to stay with his father to be less then 50%, and similarly if it specified that the child is to stay with his mother. However, they thought that there was a higher – almost 60% – chance that the parents would agree to divorce in the disjunctive case, when there was uncertainty about whether the settlement would specify that the child is to stay with the father or with the mother ($z = 4.54$ and 2.52 for the father and for the mother, respectively; $p < .05$ in both cases.) Because the above effect is small, and there are potential ambiguities in the interpretation of the problem, more exploration of this kind of judgement is

required. It does appear, however, that people's reasoning through this disjunctive situation may be nonconsequential. In effect, the pattern above may capture a "disjunction effect" in judgement similar to that previously observed in choice. Either disjunct – each branch of the tree – leads to attribute a probability of less than one-half, but when facing the disjunction people estimate a probability greater than one-half.

Disjunction effects in judgement are likely to arise in contexts similar to those which characterize these effects in choice. While either disjunct presents a clear scenario with compelling reasons for increasing or decreasing one's probability estimate, a disjunctive situation can be less compelling. Thus, people tend to suspend judgement in disjunctive situations, even if every disjunct would eventually affect their perceived likelihood is similar ways. As in choice, instead of contemplating the consequences of traversing each of the branches, people tend to remain nonconsequential at the uncertain node. In the above divorce scenario, it appears, people see a clear reason for lowering the probability estimate of a settlement once they know that the child is to stay with his father, namely, the mother is likely to object. Similarly, if the child is to stay with his mother, the father will object. But what about when the fate of the child is not known? Rather than consider the potential objections of each parent, subjects evaluate the situation from a disjunctive perspective, wherein neither parent has reasons to object. From this perspective the couple seems ready for divorce.

Of course, people do not always refrain from considering the potential implications of disjunctive inferences. The pilot data above illustrate one kind of situation that may yield such effects due to the way uncertainty renders certain considerations less compelling. More generally, such patterns can emerge from a tendency towards "concrete thinking" (Slovic, 1972) wherein people rely heavily on information that is explicitly available, at the expense of other information which remains implicit. Numerous studies have shown that people often do not decompose categories into their relevant subcategories. For example, having been told that "robins have an ulnar artery", subjects rate it more likely that all birds have an ulnar artery than that ostriches have it (Osherson, Smith, Wilkie, Lopez, & Shafir, 1990; see also Shafir, Smith, & Osherson, 1990). A precondition for such judgement is the failure to take account of the fact that the category *birds* consists of subcategories, like *robins*, *sparrows*, and *ostriches*. Along similar lines, most subjects estimate the frequency, on a typical page, of seven-letter words that end in *ing* (----*ing*) to be greater than the frequency of seven-letter words that have the letter *n* in the sixth position (----*n*-) (Tversky & Kahneman, 1983). When making these estimates, subjects focus on the particular category under consideration: because instances of the former category are more easily available than instances of the latter, subjects erroneously conclude that they must be more frequent. Evidently, subjects do not decompose the latter category into its constituent subcategories (i.e., seven-letter words that end in *ing*, seven-letter words that end in *ent*, seven-letter words that end in *ine*, etc.).

Various manifestations of the tendency for considerations that are out of sight to be out of mind have been documented by Fischhoff, Slovic, and Lichtenstein (1978) who, for example, asked car mechanics to assess the probabilities of different causes of a car's failure to start. The mean probability assigned to the hypothesis "the cause of failure is something other than the battery, the fuel system, or the engine" doubled when the unspecified disjunctive alternative was broken up into some of its specific disjuncts (e.g., the starting system, the ignition system, etc.) Along similar lines, Johnson, Hershey, Meszaros, and Kunreuther (1993) found that subjects were willing to pay more when offered health insurance that covers hospitalization for "any disease or accident" than when offered health insurance that covers hospitalization "for any reason". Evidently, subjects do not perceive the latter, implicit disjunction as encompassing the various disjuncts explicitly mentioned in the former. For an extensive treatment of the relationship between explicit and implicit disjunctions in probability judgement, see Tversky and Koehler (1993).

Inferential disjunction effects may also occur in situations in which different rationales apply to the various disjuncts. Under uncertainty, people may be reluctant to contemplate the consequences, even if they would eventually affect judgement in similar ways. Shafir and Tversky (1992) have suggested that the financial markets' behavior during the 1988 US Presidential election had all the makings of a disjunction effect. In the weeks preceding the election, US financial markets remained relatively inactive and stable, "because of caution before the Presidential election" (*The New York Times*, November 5, 1988). "Investors were reluctant to make major moves early in a week full of economic uncertainty and seven days away from the Presidential election" (*The Wall Street Journal*, November 2, 1988). Immediately following the election, a clear outlook emerged. The dollar plunged sharply to its lowest level in 10 months, stock and bond prices declined, and the Dow Jones industrial average fell a total of almost 78 points over the ensuing week. The dollar's decline, explained the analysts, "reflected continued worry about the US trade and budget deficits", "economic reality has set back in" (*WSJ*, November 10). The financial markets, observed the *NYT*, "had generally favored the election of Mr. Bush and had expected his victory, but in the three days since the election they have registered their concern about where he goes from here". Of course, the financial markets were likely to have registered at least as much concern had Mr. Dukakis been elected. Most traders agree, wrote the *WSJ*, that the stock market would have dropped significantly had Dukakis staged a come-from-behind victory. "When I walked in and looked at the screen", explained one trader after the election, "I thought Dukakis had won" (*NYT*, November 10).

After days of inactivity preceding the election, the market declined immediately following Bush's victory, and would have declined at least as much had Dukakis been the victor (these would unlikely be due to disjoint sets of actors). Of course, a thorough analysis of the financial markets' behavior reveals

numerous complications but, at least on the surface, this incident has all the makings of a disjunction effect: the markets would decline if Bush was elected, they would decline if Dukakis was elected, but they resisted any change until after the elections. Being at the node of such a momentous disjunction seems to have stopped Wall Street from addressing the expected consequences. "Considering how Wall Street had rooted for Bush's election", said the *NYT* (November 11), "its reaction to his victory was hardly celebratory. Stocks fell, bonds fell and the dollar dropped. It makes one think of the woman in the New Yorker cartoon discussing a friend's failing marriage: 'She got what she wanted, but it wasn't what she expected'." Indeed, it is in the nature of nonconsequential thinking to encounter events that were bound to be, but were not expected.

Deductive inference

The Wason selection task

One of the most investigated tasks in research into human reasoning has been the selection task, first described by Wason (1966). In a typical version of the task, subjects are presented with four cards, each of which has a letter on one side and a number on the other. Only one side of each card is displayed. For example:

Subjects' task is to indicate those cards, and only those cards, that must be turned over to test the following rule: "If there is a vowel on one side of the card, then there is an even number on the other side of the card." The simple structure of the task is deceptive – the great majority of subjects fail to solve it. Most select the *E* card or the *E* and the *4* cards, whereas the correct choices are the *E* and the *7* cards. (The success rate of initial choices in dozens of studies employing the basic form of the selection task typically ranges between 0 and a little over 20%; see Evans, 1989, and Gilhooly, 1988, for reviews.) The difficulty of the Wason selection task is perplexing. Numerous variations of the task have been documented, and they generally agree that people have no trouble evaluating the relevance of the items that are hidden on the other side of each card. Wason and Johnson-Laird (1970; see also Wason, 1969) explicitly address the discrepancy between subjects' ability to evaluate the relevance of potential outcomes (i.e., to understand the truth conditions of the rule), and their inappropriate selection of the relevant cards. (Oakhill & Johnson-Laird, 1985, report related findings regarding subjects' selection of counterexamples when testing generalizations.) While the problem is logically quite simple, they conclude, "clearly it is the attempt to solve it which makes it difficult" (Wason & Johnson-Laird, 1972, p. 174). Thus, subjects understand that neither a vowel nor a consonant on the other

side of the *4* card contributes to the possible falsification of the rule, yet they choose to turn the *4* card when its other side is not known. Similarly, subjects understand that a consonant on the other side of the *7* card would not falsify the rule but that a vowel *would* falsify it, nevertheless they neglect to turn the *7* card in the disjunctive situation. As Evans (1984, p. 458) has noted, "this strongly confirms the view that card selections are not based upon any analysis of the consequences of turning the cards". Subjects are easily able to evaluate the logical consequences of potential outcomes in isolation, but they seem to act in ways that ignore these consequences when facing a disjunction.

What exactly subjects do when performing the selection task remains outside the purview of the present paper, especially considering the numerous studies that have addressed this question. In general, a pattern of content effects has been observed in a number of variations on the task (see, for example, Griggs & Cox, 1982; Johnson-Laird, Legrenzi, & Legrenzi, 1972; and Wason, 1983, for a review; although see also Manktelow & Evans, 1979). It is likely that such content effects facilitate performance on the selection task by rendering it more natural for subjects to contemplate the possible outcomes, which tend to describe familiar situations. To explain the various effects, researchers have suggested verification biases (Johnson-Laird & Wason, 1970), matching biases (Evans, 1984; Evans & Lynch, 1973), memories of domain-specific experiences (Griggs & Cox, 1982; Manktelow & Evans, 1979), pragmatic reasoning schemas (Cheng & Holyoak, 1985, 1989), selective focusing (Legrenzi, Girotto, & Johnson-Laird, 1993), as well as an innate propensity to look out for cheaters (Cosmides, 1989). What these explanations have in common is an account of performance on the selection task that does not involve disjunctive reasoning per se. Instead, people are assumed to focus on items that have been explicitly mentioned, to apply prestored knowledge structures, or to remember relevant past experiences. While most people find it trivially easy to reason logically about each isolated disjunct, the disjunction leads them to withhold such reasoning, at least when the content is not familiar. Subjects confronted with the above four-card problem fail to consider the logical consequences of turning each card, and instead remain, judgement suspended, at the disjunctive node, the cards' hidden sides not having been adequately evaluated.

The THOG problem

Another widely investigated reasoning problem whose disjunctive logic makes it difficult for most people to solve is the THOG problem (Wason & Brooks, 1979). The problem presents four designs: a black triangle, a white triangle, a black circle, and a white circle. Subjects are given an exclusive disjunction rule. They are told that the experimenter has chosen one of the shapes (triangle or

circle) and one of the colors (black or white), and that any design is a THOG if, and only if, it has either the chosen shape or color, but not both. Told that the black triangle is a THOG, subjects are asked to classify each of the remaining designs. The correct solution is that the white circle is a THOG and that the white triangle and the black circle are not. This is because the shape and color chosen by the experimenter can only be either a circle and black, or a triangle and white. In both cases the same conclusion follows: the black circle and white triangle are not THOGs and the white circle is. The majority of subjects, however, fail to follow this disjunctive logic and the most popular answer is the mirror image of the correct response. Reminiscent of the selection task, subjects appear to have no difficulty evaluating what is and what is not a THOG once they are told the particular shape and color chosen by the experimenter (Wason & Brooks, 1979). It is when they face a disjunction of possible choices that subjects appear not to work through the consequences. Uncertain about the correct shape and color, subjects fail to consider the consequences of the two options and reach a conclusion that contradicts their preferred solution given either alternative. Smyth and Clark (1986) and Girotto and Legrenzi (1993) also address the relationship between failure on the THOG problem and nonconsequential reasoning through disjunctions.

Double disjunctions

Further evidence of subjects' reluctance to think through inferential disjunctions comes from a recent study of propositional reasoning conducted by Johnson-Laird, Byrne, and Schaeken (1992; see also Johnson-Laird & Byrne, 1991, Chapter 3). These investigators presented subjects with various premises and asked them to write down what conclusion, if any, followed from those premises. They concluded that reasoning from conditional premises was easier for all subjects than reasoning from disjunctive premises. In one study subjects were presented with "double disjunctions" – two disjunctive premises such as the following:

June is in Wales or Charles is in Scotland, but not both.
Charles is in Scotland or Kate is in Ireland, but not both.

To see what follows from this double disjunction, one simply needs to assume, in turn, the separate disjuncts. If we assume that June is in Wales, then it is not the case that Charles is in Scotland and, therefore, we know that Kate is in Ireland. Similarly, if we assume that Charles is in Scotland, then it is not the case that June is in Wales or that Kate is in Ireland. It therefore follows from this double disjunction that either Charles is in Scotland or June is in Wales and Kate is in

Ireland. It is clear, once separate disjuncts are entertained, that certain conclusions follow. Yet, nearly a quarter of Johnson-Laird et al.'s subjects (ages 18–59, all working at their own pace) concluded that nothing follows, and many others erred in their reasoning, to yield a total of 21% valid conclusions. (Other kinds of disjunctions – negative and inclusive – fared worse, yielding an average of 5% valid conclusions.) As in the previous studies, if subjects are provided with relevant facts they have no trouble arriving at valid conclusions. Thus, once subjects are told that, say, June is in Wales, they have no trouble concluding that Kate is in Ireland. Similarly, they reach a valid conclusion if told that Charles is in Scotland. But when facing the disjunctive proposition, people seem to confound their epistemic uncertainty, what they may or may not know, with uncertainty about the actual consequences, the fact that one or another of the disjuncts must obtain. Presented with a disjunction of simple alternatives most subjects refrain from assuming the respective disjuncts and arrive at no valid conclusions.[4]

Puzzles and paradoxes

The impossible barber

Many well-known puzzles and semantic paradoxes have an essentially disjunctive character. Consider, for example, that famous, clean-shaven, small-village

[4] Johnson-Laird, Byrne, and Schaeken (1992) investigate these disjunctions in the context of their theory of propositional reasoning. In fact, a number of psychological theories of propositional reasoning have been advanced in recent years (e.g., Braine, Reiser, & Rumain, 1984; Osherson, 1974–6; Rips, 1983), and the relationship between reasoning about disjunctive propositions and reasoning through disjunctive situations merits further investigation. One issue that arises out of the aforementioned studies is worth mentioning. Rips (1983), in his theory of propositional reasoning which he calls ANDS, finds reason to assume certain "backward" deduction rules that are triggered only in the presence of subgoals. This leads Braine et al. to make the following observation:

> The conditionality of inferences on subgoals places ANDS on a very short leash that has counterintuitive consequences. For example, consider the following premises:
>
> *There is an F or an R*
> *If there is an F then there is an L*
> *If there is an R then there is an L*
>
> It seems intuitively obvious that there has to be an *L*. If ANDS is given the conclusion *There is an L*, then ANDS makes the deduction. But if the conclusion given is anything else (e.g., *There is an X*, or *There is not an L*), ANDS will not notice that there has to be an *L*. (1984, pp. 357–8)

The explicit availability of the premises in the example above may distinguish it from a standard disjunction effect, wherein the specific disjuncts are not explicitly considered. Apart from that, the phenomenon – that it should be "intuitively obvious" that there has to be an *L*, but may "not be noticed" – seems a good simulation of the disjunction effect.

barber who shaves all and only the village men who do not shave themselves. The description of this barber seems perfectly legitimate – one almost feels like one may know the man. Behind this description, however, lurks an important disjunction: either this barber shaves himself or he does not (which, incidentally, still seems perfectly innocent). But, of course, once we contemplate the disjuncts we realize the problem: if the barber shaves himself, then he violates the stipulation that he only shaves those who do *not* shave themselves. And if he does not shave himself, then he violates the stipulation that he shaves *all* those who don't. The impossible barber is closely related to another of Bertrand Russell's paradoxes, namely, the set paradox. The set paradox, which had a profound influence on modern mathematical thinking, concerns the set of all sets that do not contain themselves as members. (Is this set a member of itself?) The logical solution to these paradoxes is beyond the scope of the present paper (see Russell, 1943), but their "paradoxical" nature is instructive. One definition of "paradox" is "a statement that appears true but which, in fact, involves a contradiction" (Falletta, 1983). What characterizes the paradoxes above is a logical impossibility that goes undetected partly due to their underlying disjunctive nature. Unless we delve into the appropriate disjuncts (which are themselves often not trivial to identify) and contemplate their logical consequences, these impossible disjunctions appear innocuous.

Knights and knaves

Many puzzles also rely on the surprising complexity or lack of clarity that arise in simple disjunctive situations. A class of such puzzles concerns an island in which certain inhabitants called "knights" always tell the truth, and others called "knaves" always lie. Smullyan (1978) presents a variety of knight–knave puzzles, and Rips (1989) investigates the psychology of reasoning about them. Consider, for example, the following puzzle (which the reader is invited to solve before reading further):

> There are three inhabitants, A, B, and C, each of whom is a knight or a knave. Two people are said to be of the *same type* if they are both knights or both knaves. A and B make the following statements:
>
> A: B is a knave
> B: A and C are of the same type.
>
> What is C?
>
> (Smullyan, 1978, p. 22; reprinted in Rips, 1989)

We know that A must be either a knight or a knave. If A is a knight, then his statement about B must be true, so B is a knave. If B is a knave, then his

statement that A and C are of the same type is false. Hence, since we are assuming A is a knight, C must be a knave. On the other hand, suppose A is a knave. Then his statement that B is a knave is false, and B is the knight. Hence, B's statement that A and C are of the same type is true, and since A is a knave, so is C. Thus, we have shown that C is a knave regardless of whether A is a knave or a knight. While each assumption about A leads straightforwardly to a conclusion about C, the disjunctive nature of the puzzle makes it quite difficult. And the difficulties are not negligible: about 30% of Rips' subjects stopped working on a set of such problems relatively quickly and scored at less than chance accuracy (which was 5%), and the remaining subjects averaged a low solution rate of 26% of the problems answered correctly. In fact, after reviewing subjects' think-aloud protocols, Rips (1989, p. 89) concludes that "most of the subjects' difficulties involved conceptual bookkeeping rather than narrowly logical deficiencies" (although see Johnson-Laird & Byrne, 1991, for a discussion of possible difficulties involved in more complex cases). Rips proceeds to stipulate that while the required propositional (logical) rules are equally available to everyone, "subjects differ in the ease with which they hit upon a stable solution path" (p. 109). Thus, it is not the simple logical steps that seem to create the difficulties in this case, but rather the general, conceptual "solution path" required to reason through a disjunction.

Conclusion

In their seminal *Study of Thinking*, Brunger, Goodnow; and Austin (1956) observed "the dislike of and clumsiness with disjunctive concepts shown by human subjects" (p. 181). The studies reviewed above indicate that people's dislike of and clumsiness with disjunctions extend across numerous tasks and domains. While various factors may contribute to the clumsiness with disjunctions in different domains, it nonetheless appears that a consideration of people's reluctance to think through disjunctions may shed light on common difficulties experienced in reasoning and decision making under uncertainty.

Decision difficulty is sometimes attributed to emotional factors having to do with conflict and indecision. Alternatively, it can result from the sheer complexity that characterizes many decision situations. In the context of the present paper, on the other hand, STP violations were observed in a number of simple contexts of decision and reasoning that do not seem readily attributable to either emotional factors or complexity considerations. The disjunctive scenarios reviewed in this paper were quite simple, most involving just a couple of possible disjuncts. In contrast to many complicated tasks that people perform with relative ease, these problems appear computationally very simple. They serve to highlight the discrepancy between logical complexity on the one hand and psychological

difficulty on the other. Recall, for example, Rips' observation in the context of the knights/knaves problem, to the effect that most subjects' difficulties involved conceptual bookkeeping, or arriving at a stable solution path, rather than narrowly logical deficiencies (for a related discussion, see Goldman, 1993). While it is possible that subjects occasionally forget intermediate results obtained in their reasoning process, subjects in these experiments were allowed to write things down and, besides, there often were very few intermediate steps to remember. In general, subjects appear reluctant to travel through the branches of a decision tree. Indeed, numerous studies have shown that merely encouraging subjects to systematically consider the various disjuncts often allows them to avoid common errors. In the context of the THOG problem, for example, Griggs and Newstead (1982) have shown that simply spelling out for subjects the four disjunctive possibilities reliably improves their performance. Similar effects have been shown by Johnson-Laird and Byrne (1991) in the context of the double disjunctions, and by Tversky and Shafir (1992) in the context of various disjunction effects in decision problems. Merely mentioning the few possible disjuncts can hardly be considered a major facilitation from the point of view of computational or logical complexity, but it does appear to set subjects on the right solution path, namely that of systematically contemplating the decision tree's various branches.

Typically, shortcomings in reasoning are attributed to quantitative limitations of human beings as processors of information. "Hard problems" tend to be characterized by reference to the "required amount of knowledge", the "memory load", or the "size of the search space" (cf. Kotovsky, Hayes, & Simon, 1985; Kotovsky & Simon, 1990). These limitations are bound to play a critical role in many situations. As discussed by Shafir and Tversky (1992), however, such limitations are not sufficient to account for all that is difficult about thinking. In contrast to the "frame problem" (Hayes, 1973; McCarthy & Hayes, 1969), for example, which is trivial for people but exceedingly difficult for AI, the task of thinking through disjunctions is trivial for AI (which routinely implements "tree search" and "path finding" algorithms) but is apparently quite unnatural for people.

It appears that decision under uncertainty can be thought of as another domain in which subjects exhibit a reluctance to think through disjunctive situations. Thinking through an event tree requires people to assume momentarily as true something that may in fact be false. People may be reluctant to make this assumption, especially when competing alternatives (other branches of the tree) are readily available. It is apparently difficult to devote full attention to each of several branches of an event tree (cf. Slovic & Fischhoff, 1977), particularly when it is known that most will eventually prove to be false hypothetical assumptions. Often, subjects may lack the motivation to trasverse the tree simply because they assume, as is often the case, that the problem will not be resolved by separately

evaluating the branches. We usually try to formulate problems in ways that have sifted through the irrelevant disjunctions: those that are left are normally assumed to matter.

It appears that part of what may be problematic in decision under uncertainty is more fundamental than the problems typically envisioned, concerning the difficulties involved in the estimation of likelihoods and their combination with the estimated utilities of outcomes. Situations of uncertainty, it is suggested, can be thought of as disjunctive situations: one event may occur, or another. The studies above indicate that the disjunctive logic of these uncertain situations often introduces an uncertainty of its own. Thus, even in situations in which there should be no uncertainty since the same action or outcome will eventually obtain in either case, people' reluctance to think through these scenarios often creates an uncertainty that, if it were not for this reluctance, would not be felt or observed.

As with numerous other systematic behavioral errors, the fact that people routinely commit a mistake does not, of course, mean that they are not capable of realizing it once it is apparent. Many of the patterns observed above, we suggest, reflect a failure on the part of people to detect and apply the relevant principles rather than a lack of appreciation for their normative appeal (see Shafir, 1993, for related discussion). Subjects' violations of STP in a variety of decision contexts were attributed to their failure to think through the disjunctive situation. In fact, when Tversky and Shafir (1992) first asked subjects to indicate their preferred course of action under each outcome and only then to make a decision in the disjunctive condition, the majority of subjects who opted for the same option under every outcome chose that option also when the precise outcome was not known. The frequency of disjunction effects, in other words, substantially diminishes when the logic of STP is made salient. Like other normative principles of decision making, STP is generally satisfied when its application is transparent, but is sometimes violated when it is not (Tversky & Kahneman, 1986). Because it is a general "solution path" that seems to be neglected, rather than a limitation in logical or computational skill, a proficiency in thinking through uncertain situations may be something that people can improve upon through deliberate planning and introspection. Further study of people's psychology in situations of uncertainty and in other disjunctive situations is likely to improve our understanding and implementation of reasoning in general, and of the decision making process in particular.

References

Bacharach, M., & Hurley, S. (1991). Issues and advances in the foundations of decision theory. In M. Bacharach & S. Hurley (Eds.), *Foundations of decision theory: Issues and advances* (pp. 1–38). Oxford: Basil Blackwell.

Bar-Hillel, M. (1973). On the subjective probability of compound events. *Organizational Behavior and Human Performance*, *9*, 396–406.

Bastardi, A., & Shafir, E. (1994). *On the search for and misuse of useless information*. Manuscript, Princeton University.

Braine, M.D.S., Reiser, B.J., & Rumain, B. (1984). Some empirical justification for a theory of natural propositional logic. In G.H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 18, pp. 313–371). New York: Academic Press.

Bruner, J.S., Goodnow, J.J., & Austin, G.A. (1956). *A study of thinking*. New York: Wiley.

Carlson, B.W., & Yates, J.F. (1989). Disjunction errors in qualitative likelihood judgment. *Organizational Behavior and Human Decision Processes*, *44*, 368–379.

Cheng, P.W., & Holyoak, K.J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391–416.

Cheng, P.W., & Holyoak, K.J. (1989). On the natural selection of reasoning theories. *Cognition*, *33*, 285–313.

Cosmides, L. (1989). The logic of social exchange: has natural selection shaped how humans reason? *Cognition*, *31*, 187–276.

Evans, J.St.B.T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology*, *75*, 451–468.

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

Evans, J.St.B.T., & Lynch, J.S. (1973). Matching bias in the selection task. *British Journal of Psychology*, *64*, 391–397.

Falletta, N.L. (1983). *The paradoxicon*. New York: Doubleday.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978). Fault trees: sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 330–344.

Gardner, M. (1973). Free will revisited, with a mind-bending prediction paradox by William Newcomb. *Scientific American*, *229*(1), 104–108.

Gardner, M. (1974). Reflections on Newcomb's problem: a prediction and free-will dilemma. *Scientific American*, *230*(3), 102–109.

Gibbard, A., & Harper, W.L. (1978). Counterfactuals and two kinds of expected utility. In C.A. Hooker, J.J. Leach, & E.F. McClennen, (Eds.), *Foundations and applications of decision theory*, (Vol. 1, pp. 125–162). Dordrecht: Reidel.

Gilhooly, K.J. (1988). *Thinking: Directed, undirected, and creative*, 2nd ed. San Diego, CA: Academic Press.

Girotto, V., & Legrenzi, P. (1993). Naming the parents of the THOG: Mental representation and reasoning. *Quarterly Journal of Experimental Psychology*, forthcoming.

Goldman, A.I. (1993). *Philosophical applications of cognitive science*. Boulder, CO: Westview Press.

Griggs, R.A., & Cox, J.R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, *73*, 407–420.

Griggs, R.A., & Newstead, S.E. (1982). The role of problem structure in a deductive reasoning task. *Journal of Experimental Psychology: Language, Memory and Cognition*, *8*, 297–307.

Hammond, P. (1988). Consequentialist foundations for expected utility. *Theory and Decision*, *25*, 25–78.

Hayes, P. (1973). The frame problem and related problems in artificial intelligence. In A. Elithorn & D. Jones (Eds.), *Artificial and human thinking*. San Francisco: Jossey-Bass.

Henslin, J.M. (1976). Craps and magic. *American Journal of Sociology*, *73*, 316–330.

Hofstadter, D.R. (1983). Dilemmas for superrational thinkers, leading up to a luring lottery. *Scientific American*, June. Reprinted in D.R. Hofstadter (1985). *Metamagical themas: Questing for the essence of mind and pattern*. New York: Basic Books.

Jahoda, G. (1969). *The psychology of superstition*. Harmondsworth: Penguin Books.

Johnson, E.J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, *7*, 35–51.

Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction*. Hillsdale: Erlbaum.

Johnson-Laird, P.N., Byrne, R.M.J., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, *99*, 418–439.

Johnson-Laird, P.N., Legrenzi, P., & Legrenzi, S.M. (1972). Reasoning and a sense of reality. *British Journal of Psychology*, *63*, 395–400.

Johnson-Laird, P.N., & Wason, P.C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, *1*, 134–148.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. *Cognition*, *11*, 123–141.

Kotovsky, K., Hayes, J.R., & Simon, H.A. (1985). Why are some problems hard? Evidence from tower of Hanoi. *Cognitive Psychology*, *17*, 248–294.

Kotovsky, K., & Simon, H.A. (1990). What makes some problems really hard: Explorations in the problem space of difficulty. *Cognitive Psychology*, *22*, 143–183.

Legrenzi, P., Girotto, V., & Johnson-Laird, P.N. (1993). Focussing in reasoning and decision making. *Cognition*, *49*, 37–66.

Levi, I. (1991). Consequentialism and sequential choice. In M. Bacharach & S. Hurley (Eds.), *Foundations of decision theory: Issues and advances* (pp. 92–122). Oxford: Basil Blackwell.

Manktelow, K.I., & Evans, J.St. B.T. (1979). Facilitation of reasoning by realism: Effect or non-effect? *British Journal of Psychology*, *70*, 477–488.

McCarthy, J., & Hayes, P. (1969). Some philosophical problems from the standpoint of Artificial Intelligence. In B. Meltzer & D. Michie (Eds.), *Machine intelligence*. New York: American Elsevier.

McClennen, E.F. (1983). Sure-thing doubts. In B.P. Stigum & F. Wenstop (Eds.), *Foundations of utility and risk theory with applications* (pp. 117–136). Dordrecht: Reidel.

Nozick, R. (1969). Newcomb's problem and two principles of choice. In N. Rescher (Ed.), *Essays in honor of Carl G. Hempel*. Dordrecht: Reidel.

Oakhill, J.V., & Johnson-Laird, P.N. (1985). Rationality, memory and the search for counterexamples. *Cognition*, *20*, 79–94.

Osherson, D.N. (1974–6). *Logical abilities in Children* (Vol. 2–4). Hillsdale, NJ: Erlbaum.

Osherson, D.N., Smith, E.E., Wilke, A., Lopez, A., & Shafir, E. (1990). Category based induction. *Psychological Review*, *97*, 185–200.

Quattrone, G.A., & Tversky, A. (1984). Causal versus diagnostic contingencies: On self-deception and on the voter's illusion. *Journal of Personality and Social Psychology*, *46*, 237–248.

Rapoport, A., & Chammah, A. (1965). *Prisoner's dilemma*. Ann Arbor: University of Michigan Press.

Rips, L.J. (1983). Cognitive processes in propositional reasoning. *Psychological Review*, *90*, 38–71.

Rips, L.J. (1989). The psychology of knights and knaves. *Cognition*, *31*, 85–116.

Rothbart, M., & Snyder, M. (1970). Confidence in the prediction and postdiction of an uncertain event. *Canadian Journal of Behavioral Science*, *2*, 38–43.

Russell, B. (1943). *The principles of mathematics*. 2nd ed. New York: Norton.

Savage, L.J. (1954). *The foundations of statistics*. New York: Wiley.

Shafer, [redacted] & Pearl, J. (Eds.) (1990). *Readings in uncertain reasoning*. San Mateo, CA: Morgan-Kaufmann.

Shafir, E. (1993). Intuitions about rationality and cognition. In K.I. Manktelow & D.E. Over (Eds.), *Rationality: Psychological and philosophical perspectives* (pp. 260–283). New York: Routledge.

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*, 11–36.

Shafir, E., Smith, E.E., & Osherson, D.N. (1990). Typicality and reasoning fallacies. *Memory and Cognition*, *18*, 229–239.

Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology*, *24*, 449–474.

Slovic, P. (1972). From Shakespeare to Simon: speculations – and some evidence – about man's ability to process information. *Orgeon Research Institute Research Monograph*, *12*(2).

Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 544–551.

Smullyan, R.M. (1978). *What is the name of this book? The riddle of Dracula and other logical puzzles*. New York: Simon & Schuster.

Smyth, M.M., & Clark, S.E. (1986). My half-sister is a THOG: Stragetic processes in a reasoning task. *British Journal of Psychology, 77*, 275–287.

Strickland, L.H., Lewicki, R.J., & Katz, A.M. (1966). Temporal orientation and perceived control as determinants of risk-taking. *Journal of Experimental Social Psychology, 2*, 143–151.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293–315.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business, 59*, 251–278.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*, 297–323.

Tversky, A., & Koehler, D.J. (1993). *Support theory: A nonextensional representation of subjective probability*. Manuscript, Stanford University.

Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological Science, 3*, 305–309.

Wason, P.C. (1966). Reasoning. In B.M. Foss (Ed.), *New horizons in psychology* (Vol. 1). Harmandsworth: Penguin.

Wason, P.C. (1969). Structural simplicity and psychological complexity: Some thoughts on a novel problem. *Bulletin of the British Psychological Society, 22*, 281–284.

Wason, P.C. (1983). Realism and rationality in the selection task. In J.St.B.T. Evans (Ed.), *Thinking and reasoning: Psychological approaches*. London: Routledge & Kegan Paul.

Wason, P.C., & Brooks, P.G. (1979). THOG: The anatomy of a problem. *Psychological Research, 41*, 79–90.

Wason, P.C., & Johnson-Laird, P.N. (1970). A conflict between selecting and evaluating information in an inferential task. *British Journal of Psychology, 61*, 509–515.

Wason, P.C., & Johnson-Laird, P.N. (1972). *Psychology of reasoning: Structure and content*. Cambridge, MA: Harvard University Press.