THINKING AND REASONING, 1998, *4* (4), 289–317

Individual Differences in Framing and Conjunction Effects

Keith E. Stanovich
University of Toronto, Canada
Richard F. West
James Madison University, USA

Individual differences on a variety of framing and conjunction problems were examined in light of Slovic and Tversky's (1974) understanding/acceptance principle—that more reflective and skilled reasoners are more likely to affirm the axioms that define normative reasoning and to endorse the task construals of informed experts. The predictions derived from the principle were confirmed for the much discussed framing effect in the Disease Problem and for the conjunction fallacy on the Linda Problem. Subjects of higher cognitive ability were disproportionately likely to avoid each fallacy. Other framing problems produced much more modest levels of empirical support. It is conjectured that the varying patterns of individual differences are best explained by two-process theories of reasoning (e.g. Evans, 1984, 1996; Sloman, 1996) conjoined with the assumption that the two processes differentially reflect interactional and analytic intelligence.

INTRODUCTION

A main theme of the so-called heuristics and biases literature of the 1970s and early 1980s was that human responses deviated from the response deemed normative according to various models of decision making and rational judgement, such as expected utility theory or the probability calculus (see Arkes & Hammond, 1986; Kahneman, Slovic, & Tversky, 1982). However, the theoretical interpretation of these empirical demonstrations of a gap between descriptive models and normative models has been enormously contentious

Requests for reprints should be sent to Keith E. Stanovich, Department of Human Development and Applied Psychology, Ontario Institute for Studies in Education, University of Toronto, 252 Bloor St. West, Toronto, Ontario, Canada M5S 1V6. Email: KStanovich@oise.utoronto.ca

This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada to Keith E. Stanovich and a James Madison University Program Faculty Assistance Grant to Richard F. West. The authors thank David Over and two anonymous reviewers for comments on an earlier version of the manuscript, and Walter Sa and Maggie Toplak for assistance in data coding.

(Baron, 1994; Cohen, 1981, 1983; Evans & Over, 1996; Gigerenzer, 1996; Kahneman, 1981; Kahneman & Tversky, 1983, 1996; Koehler, 1996; Stanovich, in press; Stein, 1996). For example, the gap between the normative and the descriptive can be interpreted as indicating systematic irrationalities in human cognition. Alternatively, it can be argued that the gap is due to the application of an inappropriate normative model or due to an alternative construal of the task on the part of the subject (see Cohen, 1981 and Stein, 1996 for extensive discussions of these possibilities).

Even the simplest principles of normative rationality have been the subject of intense dispute. Take, for example, the basic principle of descriptive invariance (Kahneman & Tversky, 1984, p.343) "that the preference order between prospects should not depend on the manner in which they are described." There is now a large literature on whether people do display framing effects that can be unambiguously interpreted as violations of this principle. For example, the Disease Problem of Tversky and Kahneman (1981, p.453) has been the subject of much contention:

> Problem 1. Imagine that the U.S. is preparing for the outbreak of an unusual disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows: If Program A is adopted, 200 people will be saved. If Program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved. Which of the two programs would you favor, Program A or Program B?
>
> Problem 2. Imagine that the U.S. is preparing for the outbreak of an unusual disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows: If Program C is adopted, 400 people will die. If Program D is adopted, there is a one-third probability that nobody will die and a two-thirds probability that 600 people will die. Which of the two programs would you favor, Program C or Program D?

Many subjects select alternatives A and D in these two problems, despite the fact that the two problems are redescriptions of each other and that Program A maps to Program C rather than D. This response pattern seemingly violates descriptive invariance. However, Berkeley and Humphreys (1982) argue that Programs A and C might not be descriptively invariant in subjects' interpretations. They argue that the wording of the outcome of Program A ("will be saved"), combined with the fact that its outcome is seemingly not described in the same exhaustive way as the consequences for Program B, suggests the possibility of human agency in the future which might enable the saving of *more* lives (see also, Kuhberger, 1995). The wording of the outcome of Program C ("will die") does not suggest the possibility of future human agency working to save more lives (indeed, the possibility of *losing* a few more might be inferred by some people). Under such a construal of the problem, it is no longer non-normative to choose Programs A and D. Likewise, Macdonald (1986, p.24) argues that, regarding the "200 people will be saved" phrasing, "it is unnatural to predict an exact number of cases" and that "ordinary language reads 'or more' into the interpretation of the statement."

Similarly, Jou, Shanteau, and Harris (1996) have argued that the Disease Problem's assumed underlying formula (Total Expected Loss − Number Saved = Resulting Loss) is without rationale and may be pragmatically odd for various reasons. For example, they argue (p.3) that:

> the deaths could be construed as occurring immediately after the decision to save 200 lives, or at some indefinite time in the future. If the deaths were construed as occurring at some unknown future time, they would not likely be seen as a consequence of saving 200 lives. Hence saving the lives will not be conceived as entailing the death of 400 people.

Similar debates have been spawned by claims that people violate the independence axiom of utility theory (Allais, 1953; Bell, 1982; Loomes & Sugden, 1982; Schick, 1987; Slovic & Tversky, 1974; Tversky, 1975). Whether or not subjects display the so-called conjunction fallacy in probabilistic reasoning has likewise proven controversial (Adler, 1991; Bar-Hillel, 1991; Dulany & Hilton, 1991; Fiedler, 1988; Politzer & Noveck, 1991; Tversky & Kahneman, 1983; Wolford, Taylor, & Beck, 1990). Analogous controversies surround the use of base rates (e.g. Koehler, 1996), confirmation bias (Klayman & Ha, 1987, 1989), belief bias (e.g. Evans, Over, & Manktelow, 1993), probability calibration (e.g. Keren, 1997), selection task choices (e.g. Oaksford & Chater, 1994, 1996), and many other tasks in the literature in which human performance seems to depart from normative models (for summaries of the large literature, see Baron, 1994; Evans, 1989; Evans & Over, 1996; Newstead & Evans, 1995; Osherson, 1995; Piattelli-Palmarini, 1994; Plous, 1993; Shafir & Tversky, 1995).

What most of the disputants in these controversies seem to have ignored is that—although the *modal* person in these experiments might well display an overconfidence effect, underutilise base rates, choose P and Q in the selection task, commit the conjunction fallacy, etc.—on each of these tasks, *some people give the standard normative response*. For example, in knowledge calibration studies, although the mean performance level of the entire sample may be represented by a calibration curve that indicates overconfidence, a few people do display near perfect calibration (Stanovich & West, 1998). As another example, consider the problems that the Nisbett group (e.g. Fong, Krantz, & Nisbett, 1986) have used to assess statistical thinking in everyday situations. Although the majority of people often ignore the more diagnostic but pallid statistical evidence, some actually do rely on the statistical evidence rather than the vivid case evidence (Stanovich & West, 1998). A few people even respond with P and not-Q on the notoriously difficult abstract selection task (Evans, Newstead, & Byrne, 1993).

The debates about how to interpret the descriptive/normative gap usually ignore these individual differences. Contending arguments are framed in terms of changes in modal or mean performance in response to the manipulation of variables that are purported to differentiate between explanations of the gap. For example, arguments about whether overconfidence in knowledge calibration is eliminated when a representative sampling of items is used have focused on changes in the mean overconfidence bias score (Gigerenzer, Hoffrage, & Kleinbolting, 1991; Griffin & Tversky, 1992; Juslin, Olsson, & Bjorkman, 1997). It will be argued here that such analyses need to be supplemented with a concern for individual differences[1], because the nature of individual differences and their patterns of covariance might have implications for the debates about how to interpret discrepancies between normative models and descriptive models of human behaviour[2].

THE UNDERSTANDING/ACCEPTANCE PRINCIPLE

In a 1974 article, Slovic and Tversky presented a "mock" debate between Allais and Savage about the independence axiom of utility theory. This axiom states that "if the option chosen does not affect the outcome in some states of the world, then we can ignore the ... outcomes in those states" (Baron, 1993, p.50; see Allais, 1953; Luce & Krantz, 1971; Savage, 1954). Slovic and Tversky (1974) speculated that the more the independence axiom of utility theory was understood, the more it would be accepted ("the deeper the understanding of the axiom, the greater the readiness to accept it" pp.372–373). Their argument was essentially that descriptive facts about argument endorsement should condition our inductive inferences about why human performance deviates from normative models[3]. Slovic and Tversky (1974) argued that understanding/acceptance

---

[1]See Stankov and Crawford (1996) and West and Stanovich (1997) for indications of how analyses of individual differences might have implications for interpretations of the psychological mechanism underlying the overconfidence effect.

[2]For exceptions to the general neglect of individual differences in the literature, see Jepson, Krantz, and Nisbett (1983), Roberts (1993), Slugoski, Shields, and Dawson (1993), Slugoski and Wilson (in press), and Yates, Lee, & Shinotsuka (1996).

[3]Their argument is one in a long tradition that allows descriptive facts to affect judgements of normative appropriateness. For example, Slovic (1995, p.370) refers to the "deep interplay between descriptive phenomena and normative principles." As Larrick, Nisbett, and Morgan (1993,p.332) have argued, "There is also a tradition of justifying, and amending, normative models in response to empirical considerations." Thagard and Nisbett (1983, p.265) refer to this tradition when arguing that "discovery of discrepancies between inferential behavior and normative standards may in some cases signal a need for revision of the normative standards, and the descriptions of behavior may be directly relevant to what revisions are made" see also Kyburg, 1983, 1991; March, 1988; Shafer, 1988). The assumptions underlying the naturalistic project in epistemology (e.g. Kornblith, 1985, 1993) have the same implication—that findings about how humans form and alter beliefs should have a bearing on normative theories of belief acquisition.

congruence, were it to occur, would increase our confidence in the normative appropriateness of the axioms. We might call Slovic and Tversky's argument the understanding/acceptance principle—that more reflective and engaged reasoners will be more likely to affirm the axioms that define normative reasoning.

In the present series of studies we employed this principle by examining an individual difference variable that should be a direct correlate of task understanding—the cognitive ability of the subject. Larrick et al. (1993, p.333) presented a parallel argument in their analysis of what justified the cost–benefit reasoning of microeconomics:

> Intelligent people would be more likely to use cost–benefit reasoning. Because intelligence is generally regarded as being the set of psychological properties that makes for effectiveness across environments … intelligent people should be more likely to use the most effective reasoning strategies than should less intelligent people.

In the present study, we extended the understanding/acceptance principle beyond the realm of the utility theory axioms into the debates about alternative task construals. Disagreements about appropriate problem construals in large part account for differing views about whether human cognition in fact violates descriptive invariance as well as other principles such as the conjunction rule of probability theory. Parallelling the quote from Larrick et al. (1993) just given, it is argued here that we may want to condition our inferences about rational problem construals based not only on what response the majority of people make, but also on what response the most cognitively competent subjects make. That is, we propose to turn the understanding/acceptance principle into an individual differences prediction (as did Larrick et al., 1993) and hypothesise that those individuals with cognitive/personality characteristics more conducive to deeper understanding will be more accepting of the axioms of instrumental rationality and of the problem construals of expert reasoners. Consistent with this individual differences prediction, Smith and Levin (1996) have found that two different framing effects (one analogous to the Disease Problem) were smaller among individuals higher in need for cognition—a dispositional variable associated with thoughtful analysis, deep reflection, and greater information search (Cacioppo, Petty, Feinstein, & Jarvis, 1996).

Here, we examine a variety of framing effects of the type discussed earlier. It has already been demonstrated that being forced to take more time or to provide a rationale for selections reduces framing effects (Miller & Fagley, 1991; Sieck & Yates, 1997; Takemura, 1992, 1993, 1994) and, further, that there are consistent individual differences across a variety of framing problems when a within-subject design is employed (Frisch, 1993). In the present studies, we investigate whether such individual differences covary with measures of cognitive ability in ways predicted by the understanding/acceptance principle.

Following an examination of a variety of framing problems, we examine another effect that has spawned many arguments about the normative appropriateness of various task construals—the so-called conjunction fallacy (Tversky & Kahneman, 1983).

In the studies that follow, cognitive ability was operationalised by an academic aptitude measure (the Scholastic Aptitude Test, SAT) that loads highly on psychometric g—that is, general intelligence. Matarazzo (1972) views the SAT primarily as a measure of general intelligence. Carroll (1993) concurs but suggests that the test is weighted toward crystallised intelligence in the context of the psychometric theory of fluid–crystallised intelligence (Horn & Cattell, 1967; Horn & Hofer, 1992). Fluid abilities are processes such as memory and reasoning which operate across a range of domains, whereas crystallised abilities are thought to "reflect one's experiential history, and are assessed by tests of vocabulary, general information, and nearly all types of acquired knowledge" (Salthouse, 1988, p.239). More relevant for the present study is that SAT-type measures of cognitive ability have been shown to be related to measures of intellectual engagement, reflective thought, and thorough information processing (Ackerman, 1996; Ackerman & Heggestad, 1997; Baron, 1985; Carroll, 1993; Goff & Ackerman, 1992).

PARTICIPANTS AND GENERAL METHOD

The participants were 295 undergraduate students (108 males and 187 females) recruited through an introductory psychology subject pool at a medium-sized state university. Their mean age was 19.0 years (SD = 1.3). For each problem described later, a few subjects failed to respond to one or the other version of the problem. These subjects were eliminated from the analyses for that problem, and thus the reported sample sizes for each of the problems are lower than 295.

Students were asked to indicate their verbal and mathematical SAT scores on a demographics sheet. The mean reported verbal SAT score (SAT-V) of the students was 524 (*SD* = 71); the mean reported mathematical SAT score (SAT-M) was 580 (*SD* = 78); and mean total SAT score was 1104 (*SD* = 112). These reported scores match the averages of this institution (520, 587, and 1107) quite closely (Straughn & Straughn, 1995).

The Scholastic Aptitude Test is a three-hour paper-and-pencil exam used for university admissions testing. The verbal section of the SAT test contains four types of items: antonyms, reading comprehension, verbal analogies, and sentence completion items in which the examinee chooses words or phrases to fill in a blank or blanks in a sentence. According to Carroll (1993, p.705), the mathematical section contains "varied items chiefly requiring quantitative reasoning and inductive ability." The standardised scores on the verbal and mathematical sections are added together to form the total score. In the entire population of test-takers throughout the previous two decades, total scores have averaged approximately 950 with a standard deviation of approximately 150 (Willingham, Lewis, Morgan, & Ramist, 1990). Thus the scores of the students matriculating to this institution are roughly one standard deviation above the mean of all of the prospective university students taking the test.

Participants completed the framing problem pairs (that is, framing was a within-subject variable) during a single two-hour session, in which they also completed some other tasks not part of the present investigation. They were tested in small groups of 3–4 individuals. The problems were interspersed between other unrelated tasks. For several problems, the positively and negatively framed items were maximally separated and counterbalanced; in others the items were presented adjacently. These differences will be noted as the individual problem pairs are discussed.

THE DISEASE PROBLEM

The Disease Problem of Tversky and Kahneman (1981) was presented as described earlier except that programs C and D were named the omega and epsilon programs, respectively. A total of 292 subjects completed both versions of the problem, with 148 subjects completing the positively framed version first (termed Order 1) and 144 subjects completing the negatively framed version first (termed Order 2). The two problems were separated by several unrelated tasks.

On a between-subjects basis (that is, considering the first problem received by both groups), an overall framing effect was demonstrated, with 67.6% of the subjects making the risk-averse choice in the positive frame and only 34.7% making the risk-averse choice in the negative frame. On a within-subject basis, framing effects of roughly equal magnitudes were observed for the subjects in both order conditions. In Order 1, 67.6% of the subjects were risk-averse in the positive frame and 45.9% were risk-averse in the negative frame. In Order 2, 51.4% of the subjects were risk-averse in the positive frame and 34.7% were risk-averse in the negative frame.

However, an analysis of response patterns among individual subjects also converged with earlier findings in indicating that only a minority of subjects demonstrated framing effects (Frisch, 1993; Schneider, 1992). Across both task orders, 202 of the 292 subjects were consistent on both trials (101 were consistently risk-averse and 101 were consistently risk-seeking). That 69.2% of the subjects responded consistently in a within-subject administration of this problem is consistent with the 63.6% figure obtained by Frisch (1993). The proportions consistently risk-averse (34.6%) and consistently risk-seeking (34.6%) were similar to those obtained in the Frisch study (30.3% and 33.3%, respectively). 25.0% of the sample (73 subjects) displayed a framing effect (the figure was 29.3% in Frisch's smaller sample). Finally, 5.8% of the sample (17 subjects) displayed reverse framing effects (risk-averse responses in the negative

TABLE 1
Mean SAT Scores for All Five Problems

	Response Patterns			
Problem	Framing Effect	Consistent	Reverse Framing	*F* ratio
Disease Problem	1075[a] (73)	1115[b] (202)	1098 (17)	3.52*
Coin Flip Problem	1115 (48)	1102 (194)	1101 (52)	0.28
Savings Problem	1112 (85)	1100 (199)	— (7)	0.73
Tennis Problem	1086 (113)	1112 (129)	1121 (45)	2.38
Movie Problem	1098 (164)	1112 (127)	— (2)	1.12

Mean SAT scores as a function of pattern of responding on various framing problems (number of subjects in parentheses).

df = 2, 289 for Disease Problem; 2, 291 for the Coin Flip Problem; 1, 282 for the Savings Problem; 2, 284 for the Tennis Problem; 1, 289 for the Movie Problem.

* = $P < .05$.

[a], [b] = means with different superscripts are significantly different (Scheffe post hoc).

frame and risk-seeking responses in the positive frame). This again is similar to the 7.1% figure in Frisch's (1993) study.

Table 1 presents the mean SAT scores of the subjects as a function of the pattern of responding on the Disease Problem. Because order did not interact with response pattern, the results have been collapsed across the two task orders. As is apparent from Table 1, the subjects giving a consistent response on both problems had significantly higher SAT scores (1115) than those subjects displaying a framing effect (1075). The small number of subjects displaying a reverse framing effect had mean SAT scores that were intermediate between the other two groups (1098). Thus, on the Disease Problem, the framing effect is a minority phenomenon (see Frisch, 1993; Schneider, 1992) and it is disproportionately displayed by those lower in cognitive ability. The effect size of the difference between the framing effect and consistent groups was .359 (throughout this article, effect sizes will be assessed using Cohen's *d*—a standardised measure of the difference between means in units of the pooled estimate of the standard deviation; see Rosenthal & Rosnow, 1991, pp. 302–303).

THE COIN FLIP PROBLEM

The Coin Flip problem was modelled after Problem 11 in Kahneman and Tversky (1979). The positive frame was:

> Assume that you have just been given a gift of $1000. You must now choose between two alternatives:
> a. taking an additional $500 for sure

b. flipping a coin and winning another $1000 if heads comes up or getting no additional money if tails comes up

The negative frame was:

> Assume that you have just been given a gift of $2000. But you now are forced to choose between the following two alternatives:
> a. losing $500 for sure
> b. flipping a coin and losing $1000 if heads comes up or losing nothing if tails comes up

A total of 294 subjects completed this problem, with 144 subjects completing the positively framed version first (termed Order 1) and 150 subjects completing the negatively framed version first (termed Order 2). The problems were again separated by several unrelated tasks. On a between-subjects basis (that is, considering the first problem received by both groups), the overall framing effect was small and not statistically significant. In line with the previous research showing that framing/reflection effects are extremely variable with problems of this type (Cohen, Jaffray, & Said, 1987; Fagley, 1993; Hershey & Schoemaker, 1980), 68.1% of the subjects made the risk-averse choice in the positive frame, whereas a majority in the negative frame (58.7%) did so as well.

On a within-subject basis, framing effects were either small or nonexistent across the two order conditions. In Order 1, 68.1% of the subjects were risk-averse in the positive frame and 60.4% were risk-averse in the negative frame. In Order 2, 55.3% of the subjects were risk-averse in the positive frame and 58.7% were risk-averse in the negative frame.

An analysis of response patterns among individual subjects indicated that 194 of the 294 subjects were consistent on both trials (133 were consistently risk-averse and 61 were consistently risk-seeking). 16.3% of the sample (48 subjects) displayed the expected framing effect (risk-averse in the positive frame and risk-seeking in the negative frame), but even more (52 subjects) displayed reverse framing effects (risk-averse responses in the negative frame and risk-seeking responses in the positive frame).

Table 1 presents the mean SAT scores of the subjects as a function of the pattern of responding on the Coin Flip Problem. Because order did not interact with response pattern, the results have been collapsed across the two task orders. There were no significant differences in cognitive ability among the three groups. The groups showing framing and reverse framing effects were as high in cognitive ability as those subjects responding consistently to both problems.

These results contrast with the overall framing effect in the between-subjects analysis of the Disease Problem. Although, on an individual subjects basis, a majority of subjects responded consistently on that problem, of those responding inconsistently, most showed the expected directional effect (risk-averse in the positive frame and risk-seeking in the negative frame). In contrast, the Coin Flip Problem failed to show the expected framing effect in the between-subjects analysis, within-subject analysis, and individual response analysis.

Because the results from the Coin Flip Problem contrast with those of the Disease Problem (where subjects showing a framing effect were significantly lower in cognitive ability) it might be thought that the presence of an overall framing effect is what produces the difference in cognitive ability. For example, it might be thought that the inconsistent responses in the Coin Flip Problem are essentially error variance because framing and reverse framing effects were roughly equal. However, the analysis of the next problem illustrates that the presence of the expected framing effect is not sufficient to produce a difference in cognitive ability.

THE SAVINGS PROBLEM

The previous two problems were classified as loss/gain problems by Frisch (1993). Another type of problem that has been used to study framing effects is the so-called % versus absolute amount problem (see Frisch, 1993; Thaler, 1980; Tversky & Kahneman, 1981). The version used here was taken from Experiment 2 of Frisch (1993) and the two problems were as follows:

1. Imagine that you go to purchase a calculator for $15. The calculator salesperson informs you that the calculator you wish to buy is on sale for $10 at the other branch of the store which is ten minutes away by car. Would you drive to the other store? a. no b. yes
2. Imagine that you go to purchase a jacket for $125. The jacket salesperson informs you that the jacket you wish to buy is on sale for $120 at the other branch of the store which is ten minutes away by car. Would you drive to the other store? a. no b. yes

Tversky and Kahneman (1981) and Frisch (1993) found more subjects willing to make the trip to save $5 for the calculator than for the jacket, thus violating the standard analysis of consumer behaviour (Thaler, 1980) which views the two versions as equivalent choices between travelling and gaining $5 versus the status quo. In contrast, subjects seem to be responding to the fact that the percentage saving is larger in the first case.

A total of 291 subjects completed the two versions of this problem in our study, with 150 subjects completing the calculator version first (termed Order 1) and 141 subjects completing the jacket version first (termed Order 2). The problems were separated by several unrelated tasks. On a between-subjects basis (that is, considering the first problem received by both groups), an overall framing effect was demonstrated, with 71.3% of the sample willing to make the trip to save $5 on the calculator, but only 34.8% of the sample willing to make the trip to save $5 on the jacket.

On a within-subject basis, the framing effect was a little larger when subjects were presented with the calculator option first. In Order 1, 71.3% of the subjects would travel to save $5 on the calculator, but only 39.3% would travel to save $5 on the jacket. In Order 2, 56.0% of the subjects would travel to save $5 on the calculator, but only 34.8% would travel to save $5 on the jacket.

An analysis of response patterns among individual subjects indicated that 199 of the 291 subjects were consistent on both trials (101 would consistently travel in both cases and 98 would consistently not travel). 29.2% of the sample (85 subjects) displayed the expected framing effect (they would travel to save $5 on the calculator but not the jacket), and only 7 subjects were inconsistent in the opposite direction (they would travel to save $5 on the jacket but not the calculator).

Table 1 presents the mean SAT scores of the subjects as a function of the pattern of responding on the Savings Problem. Because order did not interact with response pattern, the results have been collapsed across the two task orders. There were no significant differences in cognitive ability. The group showing a framing effect had a mean SAT score that was in fact higher than that of those responding consistently on both versions; however the means were not significantly different. Unlike the case of the Disease Problem, in the Savings Problem, those who construe the choices so as to justify a response deemed inconsistent by a standard economic analysis were not disproportionately of lower cognitive ability than those giving consistent responses according to the expert interpretation.

THE TENNIS PROBLEM

Frisch (1993) examined framing as defined by the sunk cost effect[4]. That is, the alternative choices in the two versions of a problem were the same but one version represented an opportunity to honour sunk costs. Our version of the tennis problem was adapted from Frisch (1993) and Thaler (1980). The two versions were as follows:

1. Imagine you have paid $300 to join a tennis club for 6 months. During the first week of your membership, you develop tennis elbow. It is extremely painful to play tennis. Your doctor tells you that the pain will continue for about a year. Estimate the number of times you will play tennis in the next 6 months.
2. Imagine you enjoy playing tennis. One day, on the court you develop tennis elbow. It is extremely painful to play tennis. Your doctor tells you that the pain

[4]Under some classifications this sunk cost manipulation would not be termed a framing effect (D. Kahneman, personal communication, 2 September, (1997). We follow Frisch's (1993) terminology in terming this a framing effect, but whether or not it is classified as such has no bearing on the issues investigated here. See page 211 of Sieck and Yates (1997) for a discussion of "strict" and "loose" concepts of framing.

will continue for about a year. Estimate the number of times you will play tennis in the next 6 months.

A total of 287 subjects completed both versions of this problem. Unlike the previous problems, the two versions of the Tennis Problem were not separated nor counterbalanced. They were presented adjacently and in the order given here. Overall, there was a large framing effect in the direction of honouring sunk costs. When people imagined that they had paid $300 to join a club, they estimated that they would have played 5.9 times in the next six months; whereas under the same conditions, had they not paid to join a club, they estimated that they would have played 4.2 times, $t(286) = 3.64$, $P < .001$.

An analysis of response patterns among individual subjects that showed that 113 of the 287 subjects indicated that they would have played more had they paid the $300 fee; 129 indicated that they would have played the same number of times in both circumstances; and 45 subjects displayed a reverse sunk cost effect (they indicated that they would have played *fewer* times had they paid the fee). Table 1 presents the mean SAT scores of the subjects as a function of their response pattern on the Tennis Problem. There were no significant differences in cognitive ability among the three groups, although the group displaying a sunk cost effect did have the lowest mean.

After responding to the two versions of the Tennis Problem, the subjects were also asked to respond to a comparison question adapted from Frisch (1993):

How do you compare Question 1 above to Question 2 above?
a. the situations in 1 and 2 are really the same
b. the situations in 1 and 2 are subjectively different
c. the situations in 1 and 2 are objectively different

For the analyses involving this question, the group showing a sunk cost and reverse sunk cost effect were combined in order to ensure a larger *N* in several of the categories and because the patterns for these two groups were very similar. Table 2 displays a contingency table which indicates that there was a significant difference in how the two groups of subjects responded to this question [$\xi^2(2) = 84.10$, $P < .001$]. Very few subjects who displayed a framing effect of some type thought that the two versions were really the same. Instead, over 50% thought that they were subjectively different and over 40% thought that they were objectively different. In sharp contrast, almost 50% of the subjects who responded identically to the two versions thought they were the same. Interestingly, however, 35.7% of the subjects who responded identically on the two versions thought that they were subjectively different.

Table 2 also presents the SAT scores of the individuals in the six different groups defined by the cross classification of framing and response to the comparison question. A 2 (framing vs consistent) × 3 (comparison question:

same, subjectively different, objectively different) analysis of variance conducted on the data indicated that there was a significant effect of framing response [$F(1, 281) = 7.17$, $P < .01$], a significant effect of comparison question response [$F(2, 281) = 4.72$, $P < .01$], but no interaction [$F(2, 281) = 2.26$, $P > .10$]. The direction of the effects were that the subjects responding consistently tended to have higher SAT scores, and the subjects who thought that the versions were subjectively different tended to be higher in cognitive ability. The highest SAT scores were obtained by the 46 subjects who responded similarly on the two versions but nevertheless thought that they were subjectively different. These subjects apparently thought that the subjective difference did not warrant a different response to the situation and it was these subjects who were the highest in cognitive ability.

THE MOVIE PROBLEM

Another sunk cost problem taken from Frisch (1993) was the Movie Problem, the two versions of which were:

1. You are staying in a hotel room on vacation. You paid $6.95 to see a movie on pay TV. After 5 minutes you are bored and the movie seems pretty bad. Would you continue to watch the movie or not?
   a. continue to watch b. turn it off
2. You are staying in a hotel room on vacation. You turn on the TV and there is a movie on. After 5 minutes you are bored and the movie seems pretty bad. Would you continue to watch the movie or not?
   a. continue to watch b. turn it off

A total of 293 subjects completed both versions of this problem. Like the Tennis Problem but unlike the other problems, the two versions of the Movie Problem were not separated or counterbalanced. They were presented adjacently and in the order given here. Overall, there was a large framing effect in the direction of honouring sunk costs; 62.5% of the sample thought they would

TABLE 2
The Tennis Problem

	Response Pattern					
Comparison Question	Framing Effect			Consistent		
Response	SAT	N	%	SAT	N	%
Same	980	6	3.8%	1103	64	49.6%
Subjectively Different	1109	86	54.4%	1128	46	35.7%
Objectively Different	1089	66	41.8%	1103	19	14.7%

Number of subjects displaying a framing effect on the tennis problem as a function of response on the comparison question and mean SAT scores of the various groups.

watch the movie if they had paid for it, whereas only 7.2% of the sample thought they would watch the movie if they had not paid for it.

An analysis of response patterns among individual subjects indicated that 164 displayed a framing effect in which sunk costs were honoured (they would watch the movie if they had paid for it but not if they had not paid for it); 127 subjects responded consistently (19 watching the movie in both cases and 108 not watching it in both cases); and 2 subjects displayed a reverse framing effect (the latter were eliminated in the analyses that follow). Table 1 presents the mean SAT scores of the subjects as a function of their response pattern on the Movie Problem. The difference between the group displaying a sunk cost effect and those responding consistently was not statistically significant, although the group displaying a sunk cost effect did have the lower mean.

As with the Tennis Problem, the subjects were also asked whether the two situations were really the same, subjectively different, or objectively different. Table 3 displays a contingency table which indicates that there was a significant difference in how the two groups of subjects responded to this question ($\xi^2(2) = 62.05$, $P < .001$]. Very few subjects who displayed a framing effect thought that the two versions were really the same. Instead, 50% thought that they were objectively different and almost 50% thought they were subjectively different. In sharp contrast, over 35% of the subjects who responded identically to the two versions thought that they were the same. Interestingly, however, 33.9% of the subjects who responded identically on the two versions thought that they were subjectively different and 30.7% thought that they were objectively different.

Table 3 also presents the SAT scores of the individuals in the six different groups defined by the cross classification of framing and response to the comparison question. A 2 (framing vs consistent) × 3 (comparison question same, subjectively different, objectively different) analysis of variance conducted on the data indicated that the effect of framing response and of comparison question response failed to reach significance; however, as in the Tennis Problem, the highest SAT scores were obtained by the 43 subjects who responded similarly on

TABLE 3
The Movie Problem

Comparison Question	Response Pattern					
	Framing Effect			Consistent		
Response	SAT	N	%	SAT	N	%
Same	1040	2	1.2%	1109	45	35.4%
Subjectively Different	1104	80	48.8%	1134	43	33.9%
Objectively Different	1094	82	50.0%	1091	39	30.7%

Number of subjects displaying a framing effect on the tennis problem as a function of response on the comparison question and mean SAT scores of the various groups.

the two versions but nevertheless thought that they were subjectively different. These subjects apparently thought that the subjective difference did not warrant a different response to the situation and it was these subjects who were the highest in cognitive ability.

CONJUNCTION FALLACIES: THE LINDA PROBLEM

Perhaps no finding in the heuristics and biases literature has been the subject of as much criticism as Tversky and Kahneman's (1983) claim to have demonstrated a conjunction fallacy in probabilistic reasoning. Most of the criticisms have focused on the issue of differential task construal and several critics have argued that there are alternative construals of the tasks that are, if anything, more rational than that which Tversky and Kahneman (1983) regard as normative (Adler, 1984, 1991; Hilton, 1995; Levinson, 1995; Macdonald & Gilhooly, 1990).

An example of the task interpretation criticism is provided by the most famous problem in this literature, the so-called Linda Problem (Tversky & Kahneman, 1983):

> Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Please rank the following statements by their probability, using 1 for the most probable and 8 for the least probable.
>
> a. Linda is a teacher in an elementary school
> b. Linda works in a bookstore and takes Yoga classes
> c. Linda is active in the feminist movement
> d. Linda is a psychiatric social worker
> e. Linda is a member of the League of Women Voters
> f. Linda is a bank teller
> g. Linda is an insurance salesperson
> h. Linda is a bank teller and is active in the feminist movement

Because alternative *h* is the conjunction of alternatives *c* and *f*, the probability of *h* cannot be higher than that of either *c* or *f*, yet 85% of the subjects in Tversky and Kahneman's (1983) study rated alternative *h* as more probable than *f*, thus displaying the conjunction fallacy.

It has been argued that there are subtle linguistic and pragmatic features of the problem that serve to block the use of the conjunction rule from probability theory, and that this response pattern should not be considered a reasoning error.

Macdonald and Gilhooly (1990, p.59) argue that it is possible that subjects will

> usually assume the questioner is asking the question because there is some reason to suppose that Linda might be a bank teller and the questioner is interested to find

> out if she is … If Linda were chosen at random from the electoral register and ‘bank teller’ was chosen at random from some list of occupations, the probability of them corresponding would be very small, certainly less than 1 in 100 … the question itself has suggested to the subjects that Linda could be a feminist bank teller. Subjects are therefore being asked to judge how likely it is that Linda is a feminist bank teller when there is some unknown reason to suppose she is, which reason has prompted the question itself.

Hilton (1995; see Dulany & Hilton, 1991) provides a similar explanation of subjects’ behaviour on the Linda Problem. Under the assumption that the detailed information given about the target means that the experimenter knows a considerable amount about Linda, then it is reasonable to think that the phrase “Linda is a bank teller” does not contain the phrase “and is not active in the feminist movement” because the experimenter already knows this to be the case. If “Linda is a bank teller” is interpreted in this way, then rating *h* as more probable than *f* no longer represents a conjunction fallacy.

Morier and Borgida (1984) point out that the presence of the unusual conjunction “Linda is a bank teller and is active in the feminist movement” itself might prompt an interpretation of “Linda is a bank teller” as “Linda is a bank teller and is not active in the feminist movement”. To avoid such an interpretation, Morier and Borgida (1984) ran a condition in which “Linda is a bank teller and is not active in the feminist movement” was included as an alternative along with “Linda is a bank teller” but this manipulation did little to reduce the conjunction fallacy. Actually, Tversky and Kahneman (1983) themselves had concerns about such an interpretation of the “Linda is a bank teller” alternative and ran a condition in which this alternative was rephrased as “Linda is a bank teller whether or not she is active in the feminist movement”. They found that conjunction fallacy was reduced from 85% of their sample to 57% when this alternative was used. Macdonald and Gilhooly (1990) did observe a much larger reduction in the fallacy with the wording “Linda is a bank teller who may or may not be active in the feminist movement” (along with some other problem alterations). Several other investigators have suggested that pragmatic inferences lead to seeming violations of the logic of probability theory in the Linda Problem (see Adler, 1991; Dulany & Hilton, 1991; Politzer & Noveck, 1991). These criticisms all share the implication that actually displaying the conjunction fallacy is a rational response triggered by the adaptive use of social cues, linguistic cues, and background knowledge (see Hilton, 1995).

Again, as a context for this interpretation, we examined just who was making this interpretation in terms of cognitive ability—in short, were the subjects making the non-extensional, pragmatic interpretation those who were disproportionately of higher cognitive ability? Because this group is in fact the majority in most studies, adaptionist models of human cognition (e.g. Anderson, 1990) might be thought to predict that they would be subjects of higher computational capacity.

In the present study, we examined the performance of 150 subjects on the Linda Problem as given earlier. A within-subject version of the conjunction-judgement task, it represents what Tversky and Kahneman (1983) call a "direct-subtle" test of susceptibility to the conjunction fallacy—where the conjunction and its constituents are directly compared by the same subjects, but the inclusion relation is not emphasised.

Consistent with the results of previous experiments on this problem (Tversky & Kahneman, 1983) 80.7% of our sample (121 subjects) displayed the conjunction effect—they rated the feminist bank teller alternative as more probable than the bank teller alternative. Table 4 presents the mean SAT Total scores of those subjects who displayed the conjunction fallacy and those who did not. The 29 subjects who did not display the conjunction fallacy had significantly higher SAT scores and the difference of 82 points was quite sizeable. It translates into an effect size of .746, which Rosenthal and Rosnow (1991, p.446) classify as "large".

THE JOB PROBLEM: AN EASIER CONJUNCTION SCENARIO

Reeves and Lockhart (1993) have demonstrated that the incidence of the conjunction fallacy can be decreased if extensional reasoning is more strongly cued by using problems that describe the event categories in some finite population. The Job Problem was adapted from their paper and completed by 149 of our subjects:

> John is a student having trouble paying his tuition. To improve his financial situation, John has applied for three different part-time jobs. For the variety-store job, there are 5 other applicants; for the bookstore job, there are 7 other applicants; and for the shoe-sales job, there is only 1 other applicant. Please rank the following statements by their probability, using 1 for the most probable and 8 for the least probable. When ranking, please use each of the numbers from 1 to 8:

TABLE 4
The Three Conjunction Effect Problems

	Incorrect	Correct	t value	Effect Size[a]
Linda Problem	1080 (121)	1162 (29)	3.58**	.746
Job Problem	1072 (57)	1111 (92)	2.06*	.349
Student Problem	1075 (35)	1103 (107)	1.28	.250

Mean SAT total scores of subjects who gave the correct and incorrect responses to the three conjunction effect problems (number of subjects in parentheses).

df = 148 for the Linda Problem, 147 for the Job Problem, and 140 for the Student Problem.

* = $P < .05$, ** = $P < .01$, all two-tailed.

[a]Cohen's *d*

a. John will be offered the variety-store job
b. John will not be offered any job
c. John will be offered the shoe-sales job
d. John will be offered the variety-store job and the shoe-sales job
e. John will be offered the variety-store job or the bookstore job
f. John will be offered the bookstore job
g. John will be offered more than one job
h. John will be offered the bookstore job and the shoe-sales job

In contrast to the Linda Problem where 80.7% of the sample displayed the conjunction fallacy, only 38.3% (57 subjects) rated *d* as more probable than *a* (the least probable of the two conjuncts). The second conjunction, *h*, displayed similar results—only 34.9% displaying the conjunction fallacy—so our focus will be on conjunction *d*. Table 4 presents the mean SAT Total scores of those subjects who displayed the conjunction fallacy and those who did not. The 92 subjects who did not display the conjunction fallacy had significantly higher SAT scores, but the difference was not as large as that displayed in the Linda Problem. The effect size of .349 is between "moderate" (.50) and "small" (.20) according to Rosenthal and Rosnow (1991, p.446). Furthermore, as Table 5 indicates, the higher SAT scores of those responding correctly on the Job Problem were almost entirely due to those who *also* responded correctly on the Linda Problem. Of the 92 responding correctly on the Job Problem, the 24 who also responded correctly on the Linda Problem had mean SAT scores of 1180; whereas the 68 who responded incorrectly on the Linda Problem had mean SAT scores that were much lower (1087) and that were not significantly different from those responding incorrectly on the Job Problem as well (1071), $t(118) = .80$, $P > .10$.

THE STUDENT PROBLEM: FREQUENCY ESTIMATION

Tversky and Kahneman (1983) and Fiedler (1988) reduced the incidence of the conjunction fallacy by having subjects estimate the frequency of the categories rather than judge probabilities (see Gigerenzer, 1991, 1993). We employed the following problem (modelled on Fieldler, 1988) which encouraged the subjects to operate in frequentistic mode:

A survey of a random sample of 100 high school seniors in Columbus, Ohio, was conducted. Please give your best estimate of the following values:

a. How many of the 100 students were planning on attending a university or community college?
b. How many of the 100 students had smoked marijuana and had had intercourse?
c. How many of the 100 students had experimented with cocaine?
d. How many of the 100 students participated on interscholastic sports teams?

e. How many of the 100 students had at least a B average and were planning on attending a university or community college?
f. How many of the 100 students had a full time job lined up after graduation?
g. How many of the 100 students had at least a B average?

In contrast to the Linda Problem where 80.7% of the sample displayed the conjunction fallacy, only 24.6% (35 of 142 subjects) gave alternative *e* a higher frequency estimate than *g* (the least frequent of the two conjuncts). Table 4 presents the mean SAT Total scores of those subjects who displayed the conjunction fallacy and those who did not. The 107 subjects who did not display the conjunction fallacy had somewhat higher SAT scores than those who did display the fallacy, but the difference was not statistically significant and it was not nearly as large as that displayed in the Linda Problem. Furthermore, as Table 5 indicates, of the 107 responding correctly on the Student Problem, the 24 who also responded correctly on the Linda Problem had mean SAT scores that were quite high (1167); whereas the 83 who responded incorrectly on the Linda Problem had mean SAT scores that were much lower (1085) and that were not significantly different from those responding incorrectly on the Student Problem as well (1065), $t(112) = .88$, $P > .10$.

Table 6 displays even more clearly the pattern of cognitive ability differences on the three conjunction problems. The major pattern is easily summarised. It is correct responding on the Linda Problem specifically that is associated with higher cognitive ability. As Table 6 indicates, subjects getting either the Job or the Student Problem correct, but not the Linda Problem, had SAT scores (1076 and 1083, respectively) only modestly higher than those subjects displaying the conjunction fallacy on all three problems (1051). Subjects responding correctly to both the Job and Student Problems, but who did not respond correctly on the

TABLE 5
The Two Additional Conjunction Problems and the Linda Problem

	Linda Problem			
	Incorrect		Correct	
Job Problem Incorrect	1071	(52)	–	(5)
Job Problem Correct	1087	(68)	1180	(24)
Student Problem Incorrect	1065	(31)	–	(4)
Student Problem Correct	1085	(83)	1167	(24)

Mean SAT total scores as a function of performance on the two additional conjunction problems conditionalised on performance on the Linda Problem (number of subjects in parentheses).

TABLE 6
The Three Conjunction Effect Problems: Further Analysis

	Mean SAT Total Score
All Three Problems Incorrect	1051 (13)
Job Problem Only Correct	1076 (18)
Student Problem Only Correct	1083 (35)
Job & Student Problems Correct (Linda Problem Incorrect)	1086 (48)
All Three Problems Correct	1182 (21)

Mean SAT total scores of subjects who gave the correct and incorrect responses to the three conjunction effect problems (number of subjects in parentheses).

Linda Problem, had SAT scores (1086) barely higher than those getting only one of the former problems correct; but subjects responding correctly on those two *and* the Linda Problem had substantially higher scores (1182).

GENERAL DISCUSSION

The majority of subjects agree with the expert consensus that the two versions of the Disease Problem should be treated as descriptively invariant. The minority who assessed the two versions differently were disproportionately of lower cognitive ability. In short, the majority of respondents, majority of experts, and the untutored subjects with the greatest capacity for considered judgement all agree on a principle of rational indifference (Broome, 1990; Schick, 1987, 1997) that classifies the two versions of this problem as descriptively invariant.

On both of the sunk cost problems there were mild tendencies for the subjects displaying framing effects to have somewhat lower SAT scores, but these tendencies did not reach statistical significance in the one-way analyses (Table 1). However, in a two-way ANOVA taking into account responses on the direct comparison question, the effect of framing was statistically significant in the Tennis Problem. For both sunk cost problems, the subjects with the highest SAT scores were those who thought that the versions were subjectively different yet nonetheless responded the same to both. These may well be subjects who were conscious of having two systems of thought in conflict—an heuristic system prone to honour sunk costs because of automatically activated schemata and an analytic reasoning system prone to objective comparison and consequentialist decisions (see Baron, 1994; Epstein, 1994; Evans, 1984, 1996; Evans & Over, 1996; Sloman, 1996). That the latter system ultimately determined the response is consistent with these subjects being high in analytic cognitive resources (see later discussion).

Large cognitive ability differences were observed on the Linda Problem. Unlike the Disease Problem, where the framing effect was a minority phenomenon, the conjunction fallacy was displayed on this problem by a substantial majority of the subjects (80.7%); however, the minority who avoided the fallacy were considerably higher in cognitive ability. The SAT differences were substantially smaller on two problems (the Job Problem and the Student Problem) which contained features that reduce conjunction effects (event categories from an obviously finite population and frequency estimation, respectively). There has been little controversy over the construals of these two problems, however. In contrast, both the Disease Problem and the Linda Problem have produced challenges to the consensus opinion on how these problems should be interpreted by subjects. Several critics (e.g. Adler, 1984, 1991; Berkeley & Humphreys, 1982; Hilton, 1995; Levinson, 1995; Macchi, 1995; Macdonald & Gilhooly, 1990) have argued that rational conversational implicatures dictate construals different from those championed by Tversky and Kahneman (1981, 1983)[5] and that the expert opinion on what is a rational construal of such problems in the heuristics and biases literature should be revised. Margolis (1987, p.158) states: "many critics have insisted that in fact it is Kahneman & Tversky, not their subjects, who have failed to grasp the logic of the problem." Messer and Griggs (1993, p.195) point out that "if a 'fallacy' is involved, it is probably more attributable to the researchers than to the subjects."

Indeed, Macdonald (1986, p.15) asks "why do Tversky and Kahneman differ from the rest of the world in the answers to their problems?" As the data presented here indicate, Tversky and Kahneman's interpretation does *not* differ from "the rest of the world" on the Disease Problem. A within-subject comparison of choices on both versions of that problem indicates that only a minority of subjects display a framing effect and they are disproportionately of lower cognitive ability—just as the understanding/acceptance principle of Slovic and Tversky (1974) predicts.

Tversky and Kahneman's interpretation of the Linda Problem does differ from that of the majority of untutored subjects but, interestingly, their interpretation is endorsed by the untutored[6] subjects who were high in analytic cognitive ability. If we accept that such individuals, like those high in need for cognition (see Smith & Levin, 1996), are more likely to deeply comprehend the problem (an assumption for which there is some evidence, see Ackerman & Heggestad, 1997; Baron, 1985; Carroll, 1993) then according to Slovic and Tversky's (1974) understanding/acceptance principle, we might infer that this covariance between performance and cognitive ability further validates the standard construal of this problem.

---

[5]Kahneman and Tversky themselves (1982 pp.132–135) have discussed the issue of conversational implicatures in laboratory experiments.

[6]On a demographics sheet, the subjects indicated whether or not they had had a logic course, whether they had had a statistics course, and the extent of their mathematics training in high school and college. Few had had the two former courses, and none of these background variables predicted responses on any of the tasks.

One possible interpretation of the individual differences displayed here is in terms of two-process theories of reasoning (Epstein, 1994; Evans, 1984, 1996; Evans & Over, 1996; Sloman, 1996). For example, Sloman (1996) distinguishes an associative processing system with computational mechanisms that reflect similarity and temporal contiguity, and a rule-based system that operates on symbolic structures having logical content. The key feature that defines the existence of two systems in a reasoning situation is that of simultaneous contradictory belief; according to Sloman (1996, p.11): "a feeling or conviction that a response is appropriate even if it is not strong enough to be acted on." Certainly such a conflict can be said to be present for the 50.4% of the subjects who responded identically to both versions of the Tennis Problem despite viewing them as subjectively or objectively different (64.6% in the Movie Problem). Sloman (1996) views the Linda Problem as the quintessence of this type of situation. He quotes Stephen Gould's introspection (Gould, 1991, p.469) that "I know the [conjunction] is least probable, yet a little homunculus in my head continues to jump up and down, shouting at me—'but she can't be a bank teller; read the description'." According to Sloman (1996), the associative system responds to the similarity (representativeness in the terminology of Tversky & Kahneman, 1983) in the conjunction, whereas the rule-based system engages probabilistic concepts which dictate that bank teller is more probable. A parallel analysis could be made using the dual process theory of Evans (1996; Evans & Over, 1996) and its distinction between implicit and explicit processes "in which tacit and parallel processes of thought combine with explicit and sequential processes in determining our actions" (Evans & Over, 1996, p.143).

We conjecture here that large differences in cognitive ability will only be found on problems that strongly engage both reasoning systems and in which the reasoning systems cue opposite responses. This is because the two systems are identified with different types of intelligence. Clearly, the rule-based system embodies analytic intelligence of the type measured on SAT tests (Carpenter, Just, & Shell, 1990; Carroll, 1993). The associative system, in contrast, might be better identified with what Levinson (1995) terms interactional intelligence. He speculates that evolutionary pressures were focused more on negotiating cooperative mutual intersubjectivity than on understanding the natural world. Because he views the primary evolutionary pressures as intraspecific, he posits (1995, p.223) that there is "a systematic bias in human thinking in other domains which might be attributed to the centrality of interactional intelligence in our intellectual makeup." Having as its goals the ability to model other minds in order to read intention and to make rapid interactional moves based on those modelled intentions, interactional intelligence is composed of the mechanisms that support a Gricean theory of communication that relies on intention-attribution. These pragmatic heuristics have, according to Levinson (1995), the property of speed and also of nonmonotonicity. They are subjectively determinate rather than probabilistic.

According to Levinson (1995, p.238), the interactional intelligence behind conversational understanding operates with an important default—that the conversational puzzles it is trying to solve were "*designed* to be solved and the clues have been designed to be sufficient to yield a determinate solution." Levinson (1995) proposes that this assumption poses "spill-over" problems when interactional intelligence, rather than analytic intelligence, is used to decode nondeterminate and nondesigned problems such as theories about nature and the human body. Levinson (1995) points out the formal similarity of the properties of interactional intelligence to some of Tversky and Kahneman's (1974, 1981, 1983) biases (salience, prototypicality, representativeness).

Levinson's (1995) analysis of interactional intelligence confronting the Linda Problem is similar to that of other theorists who emphasise the importance of conversational implicatures (Adler, 1984, 1991; Hilton, 1995; Macchi, 1995; Schwarz, 1996). Under the presumption of experimenter cooperativeness, only the relevant facts would be presented in Linda's description. If the facts presented are both relevant and correct then either the "bank teller" should not be considered (as it so clearly contradicts the information given) or it must be meant by the experimenter to be interpreted as "bank teller who is not a feminist".

Using the distinction between analytic and interactional intelligence—and the distinction between processing systems articulated by Sloman (1996) and others (e.g. Evans, 1996; Evans & Over, 1996)—it is conjectured here that in order to observe large cognitive ability differences in a decision-making situation, two conditions are necessary. First, the task must engage both the associative and the rule-based system. This will happen quite often because, as Evans and Over (1996, p.144) note: "almost all reasoning tasks show evidence of a logical and non-logical component of performance." Second, these two systems must strongly cue *different* responses. It is not enough simply that both systems are engaged[7]. If both cue the same response, then this could have the effect of severely diluting any differences in cognitive ability. One reason that we predict this outcome is that we assume individual differences in interactional intelligence bear little relation to individual differences in analytic intelligence. This is a conjecture for which there is a modest amount of evidence. Individual differences in implicit associative induction have displayed much smaller correlations with analytic intelligence than have individual differences in rule-based reasoning (McGeorge, Crawford, & Kelly, 1997; Reber, 1993; Reber, Walkenfeld, & Hernstadt, 1991). Furthermore, direct indicators of interactional intelligence have displayed very low correlations with measures of analytic intelligence (Matthews & Keating, 1995).

If this conjecture is correct, then the associative system will equally cue subjects of high and low analytic intelligence. If the associative system cues a

[7]Of course, another way that cognitive ability differences might be observed is if the task engages only the rule-based system. For the present discussion, this is an uninteresting case.

response that is also signalled by the rule-based system, it will tend to dilute any cognitive ability differences by drawing to the response equally individuals high and low in analytic intelligence. In contrast, if the two systems cue opposite responses, the rule-based system will tend to differentially cue those of high analytic intelligence and this tendency will not be diluted by the associative system nondifferentially drawing subjects to the same response.

In the Linda Problem, the associative system is the dominant cueing system (80.7% showing the conjunction fallacy), whereas in the Disease Problem the rule-based system is the dominant cueing system (only 25.0% showing a framing effect). Despite the differing strengths of the two systems in these two cases, the two problems both display cognitive ability differences because in both problems the two processing systems are cueing different responses. Ability differences may be attenuated on other problems (none of which has caused the level of contentious debate as have the Linda or Disease Problems) because the two processes are not as cleanly associated with alternative responses (the rule-based system not as strongly dictating the normative response and the associative response not as strongly cueing the non-normative response). Note for example that, consistent with other research (Cohen et al., 1987; Schneider, 1992), the Coin Flip Problem displayed no overall framing effect. It may be that the rule-based system *and* the associative system (contra prospect theory) are cueing consistency in this problem, and that the relatively equal numbers of framing and reverse framing effects represent error variance around the normative response of consistency. The lack of cognitive ability differences on this problem is consistent with our conjecture that there is a lack of differential cueing.

Similarly, performance across the three conjunction problems is consistent with the differential cueing view. The Linda Problem maximises the tendency for the associative and rule-based systems to prime different responses, and this problem displayed the largest difference in cognitive ability. The other two conjunction problems removed some of the conflicts between the two systems (primarily the tendency for the associative system to cue a nonextensional response) and the cognitive ability difference decreased on these two problems. More research is needed to clarify whether differential cueing of the two processing systems can account for the patterns of cognitive ability differences observed in framing and conjunction problems.

Further research might also indicate whether our identification of analytic and interactional intelligence with the different sets of processes in the two-process theories of Sloman (1996) and Evans and Over (1996) is a useful theoretical step. Regardless of the outcome of that theoretical programme, the present studies provide a demonstration of how analyses of individual differences can be used—in conjunction with tools such as the understanding/acceptance assumption—to help explain instances where descriptive and normative models of human reasoning do not coincide. The understanding/acceptance principle may become one of a small set of empirical tools available for adjudicating disputes about the appropriateness of evaluating human performance against certain formal models of normative rationality.

Manuscript received 25 March 1997
Revised manuscript received 22 January 1998

REFERENCES

Ackerman, P.L. (1996). A theory of adult development: Process, personality, interests, and knowledge. *Intelligence* , *22*, 227–257.

Ackerman, P.L., & Heggestad, E.D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*, 219–245.

Adler, J.E. (1984). Abstraction is uncooperative. *Journal for the Theory of Social Behaviour*, *14*, 165–181.

Adler, J.E. (1991). An optimist's pessimism: Conversation and conjunctions. In E. Eells & T. Maruszewski (Eds.), *Probability and rationality: Studies on L. Jonathan Cohen's philosophy of science* (pp.251–282). Amsterdam: Editions Rodopi.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axioms de l'scole americaine. *Econometrica* , *21*, 503–546.

Anderson, J.R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Arkes, H., & Hammond, K. (Eds.) (1986). *Judgment and decision making*. Cambridge, UK: Cambridge University Press.

Bar-Hillel, M. (1991). Commentary on Wolford, Taylor, and Beck: The conjunction fallacy? *Memory & Cognition*, *19*, 412–414.

Baron, J. (1985). What kinds of intelligence components are fundamental? In S. Chipman & J. Segal (Eds.), *Thinking and Learning skills (Vol. 2)* (pp.365–390). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Baron, J. (1993). *Morality and rational choice*. Dordrecht: Kluwer.

Baron, J. (1994). Nonconsequentialist decisions. *Behavioral and Brain Sciences*, *17*, 1–42.

Bell, D.E. (1982). Regret in decision making under uncertainty. *Operations Research*, *30*, 961–981.

Berkeley, D., & Humphreys, P. (1982). Structuring decision problems and the "bias heuristic". *Acta Psychologica*, *50*, 201–252.

Broome, J. (1990). Should a rational agent maximize expected utility? In K.S. Cook & M. Levi (Eds.), *The limits of rationality* (pp.132–145). Chicago: University of Chicago Press.

Cacioppo, J.T., Petty, R.E., Feinstein, J., & Jarvis, W. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*, 197–253.

Carpenter, P.A., Just, M.A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*, 404–431.

Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Cohen, L.J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, *4*, 317–370.

Cohen, L.J. (1983). The controversy about irrationality. *Behavioral and Brain Sciences*, *6*, 510–517.

Cohen, M., Jaffray, J., & Said, T. (1987). Experimental comparison of individual behavior under risk and under uncertainty for gains and losses. *Organizational Behavior and Human Decision Processes*, 39, 1–22.

Dulany, D.E., & Hilton, D.J. (1991). Conversational implicature, conscious representation and the conjunction fallacy. *Social Cognition*, *9*, 85–110.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, *49*, 709–724.

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*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Evans, J.St.B.T. (1996). Deciding before you think: Relevance and reasoning in the selection task. *British Journal of Psychology*, *87*, 223–240.

Evans, J.St.B.T., Newstead, S.E., & Byrne, R.M.J. (1993). *Human reasoning: The Psychology of deduction*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Evans, J.St.B.T., & Over, D.E. (1996). *Rationality and reasoning*. Hove, UK: Psychology Press.

Evans, J.St.B.T., Over, D.E., & Manktelow, K. (1993). Reasoning, decision making and rationality. *Cognition*, *49*, 165–187.

Fagley, N.S. (1993). A note concerning reflection effects versus framing effects. *Psychological Bulletin*, *113*, 451–452.

Fiedler K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research*, *50*, 123–129.

Fong, G.T., Krantz, D.H., & Nisbett, R.E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, *18*, 253–292.

Frisch, D. (1993). Reasons for framing effects. *Organizational Behavior and Human Decision Processes*, *54*, 399–429.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". *European Review of Social Psychology*, *2*, 83–115.

Gigerenzer, G. (1993). The bounded rationality of probabilistic mental models. In K. Manktelow & D. Over (Eds.), *Rationality: Psychological and philosophical perspectives* (pp.284–313). London: Routledge.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychological Review*, *103*, 592–596.

Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506–528.

Goff, M., & Ackerman, P.L. (1992). Personality–intelligence relations: Assessment of typical intellectual engagement. *Journal of Educational Psychology*, *84*, 537–552.

Gould, S.J. (1991). *Bully for the Brontosaurus*. New York: Norton.

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, *24*, 411–435.

Hershey, J.C., & Schoemaker, P. (1980). Prospect theory's reflection hypothesis: A critical examination. *Organizational Behavior and Human Performance*, *25*, 395–418.

Hilton, D.J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin*, *118*, 249–271.

Horn, J.L., & Cattell, R.B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica* , *26*, 1–23.

Horn, J.L., & Hofer, S. (1992). Major abilities and development in the adult period. In R.J. Sternberg & C.A. Berg (Eds.), *Intellectual development* (pp.44–99). Cambridge: Cambridge University Press.

Jepson, C., Krantz, D., & Nisbett, R. (1983). Inductive reasoning: Competence or skill? *Behavioral and Brain Sciences*, *6*, 494–501.

Jou, J., Shanteau, J., & Harris, R.J. (1996). An information processing view of framing effects: The role of causal schemas in decision making. *Memory & Cognition*, *24*,1–15.

Juslin, P, Olsson, H., & Bjorkman, M. (1997). Brunswikian and Thurstonian origins of bias in probability assessment: On the interpretation of stochastic components of judgment. *Journal of Behavioral Decision Making*, *10*, 189–209.

Kahneman, D. (1981). Who shall be the arbiter of our intuitions? *Behavioral and Brain Sciences*, **4**, 339–340.

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

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.

Kahneman, D., & Tversky, A. (1983). Can irrationality be intelligently discussed? *Behavioral and Brain Sciences*, *6*, 509–510,

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*, 341–350.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, 103, 582–591.

Keren, G. (1997). On the calibration of probability judgments: Some critical comments and alternative perspectives. *Journal of Behavioral Decision Making*, *10*, 269–278.

Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211–228.

Klayman, J., & Ha, Y. (1989). Hypothesis testing in rule discovery: Strategy, structure, and content. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **15**, 596–604.

Koehler, J.J. (1996). The base rate fallacy reconsidered: Descriptive, normative and methodological challenges. *Behavioral and Brain Sciences*, *19*, 1–53.

Kornblith, H. (Ed.) (1985). *Naturalizing epistemology*. Cambridge, MA: MIT Press.

Kornblith, H. (1993). *Inductive inference and its natural ground*. Cambridge, MA: MIT Press.

Kuhberger, A. (1995). The framing of decisions: A new look at old problems. *Organizational Behavior and Human Decision Processes*, *62*, 230–240.

Kyburg, H.E. (1983). Rational belief. *Behavioral and Brain Sciences*, *6*, 231–273.

Kyburg, H.E. (1991). Normative and descriptive ideals. In J. Cummins & J. Pollock (Eds.), *Philosophy and AI: Essays at the interface* (pp.129–139). Cambridge, MA: MIT Press.

Larrick, R.P., Nisbett, R.E., & Morgan, J.N. (1993). Who uses the cost–benefit rules of choice? Implications for the normative status of microeconomic theory. *Organizational Behavior and Human Decision Processes*, *56*, 331–347.

Levinson, S.C. (1995). Interactional biases in human thinking. In E. Goody (Ed.), *Social intelligence and interaction* (pp.221–260). Cambridge: Cambridge University Press,

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, *92*, 805–824.

Luce, R.D., & Krantz, D.H. (1971). Conditional expected utility. *Econometrica* , *39*, 253–271.

Macchi, L. (1995). Pragmatic aspects of the base-rate fallacy. *Quarterly Journal of Experimental Psychology*, *48A*, 188–207.

Macdonald, R. (1986). Credible conceptions and implausible probabilities. *British Journal of Mathematical and Statistical Psychology*, *39*, 15–27.

Macdonald, R.R., & Gilhooly, K.J. (1990). More about Linda *or* conjunctions in context. *European Journal of Cognitive Psychology*, *2*, 57–70.

March, J.G. (1988). Bounded rationality, ambiguity, and the engineering of choice. In D. Bell, H. Raiffa, & A. Tversky (Eds.), *Decision making: Descriptive, normative, and prescriptive interactions* (pp.33–57). Cambridge: Cambridge University Press.

Margolis, H. (1987). *Patterns, thinking, and cognition*. Chicago: University of Chicago Press.

Matarazzo, J.D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th Edn.). Baltimore: The Williams & Wilkins Co.

Matthews, D.J., & Keating, D.P. (1995). Domain specificity and habits of mind. *Journal of Early Adolescence* , *15*, 319–343.

McGeorge, P., Crawford, J., & Kelly, S. (1997). The relationships between psychometric intelligence and learning in an explicit and an implicit task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 239–245.

Messer, W.S., & Griggs, R.A. (1993). Another look at Linda. *Bulletin of the Psychonomic Society*, *31*, 193–196.

Miller, P.M., & Fagley, N.S. (1991). The effects of framing, problem variations, and providing rationale on choice. *Personality and Social Psychology Bulletin*, *17*, 517–522.

Morier, D.M., & Borgida, E. (1984). The conjunction fallacy: A task specific phenomenon? *Personality and Social Psychology Bulletin*, *10*, 243–252.

Newstead, S.E., & Evans, J.St.B.T. (Eds.) (1995). *Perspectives on thinking and reasoning*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*, 608–631.

Oaksford, M., & Chater, N. (1996). Rational explanation in the selection task. *Psychological Review*, *103*, 381–391.

Osherson, D.N. (1995). Probability Judgment. In E.E. Smith & D.N. Osherson (Eds.), *Thinking (Vol. 3)* (pp.35–75). Cambridge, MA: MIT Press.

Piattelli-Palmarini, M. (1994). *Inevitable illusions: How mistakes of reason rule our minds*. New York: John Wiley.

Plous, S. (1993). *The psychology of judgment and decision making*. New York: McGraw-Hill.

Politzer, G., & Noveck, I.A. (1991). Are conjunction rule violations the result of conversational rule violations? *Journal of Psycholinguistic Research*, *20*, 83–103.

Reber, A.S. (1993). *Implicit learning and tacit knowledge*. New York: Oxford University Press.

Reber, A.S., Walkenfeld, F.F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 888–896.

Reeves, T., & Lockhart, R.S. (1993). Distributional versus singular approaches to probability and errors in probabilistic reasoning. *Journal of Experimental Psychology: General*, *122*, 207–226.

Roberts, M.J. (1993). Human reasoning: Deduction rules or mental models, or both? *Quarterly Journal of Experimental Psychology*, *46*A, 569–589.

Rosenthal, R., & Rosnow, R.L. (1991). *Essentials of behavioral research: Methods and data analysis (2nd Edn)*. New York: McGraw-Hill.

Salthouse, T.A. (1988). Resource-reduction interpretations of cognitive aging. *Developmental Review*, *8*, 238–272.

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

Schick, F. (1987). Rationality: A third dimension. *Economics and Philosophy*, *3*, 49–66.

Schick, F. (1997). *Making choices: A recasting of decision theory*. Cambridge: Cambridge University Press.

Schneider, S.L. (1992). Framing and conflict: Aspiration level contingency, the status quo, and current theories of risky choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1040–1057.

Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods, and the logic of conversation*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Shafer, G. (1988). Savage revisited. In D. Bell, H. Raiffa, & A. Tversky (Eds.), *Decision making: Descriptive, normative, and prescriptive interactions* (pp.193–234). Cambridge: Cambridge University Press.

Shafir, E., & Tversky, A. (1995). Decision making. In E.E. Smith & D.N. Osherson (Eds.), *Thinking (Vol. 3)* (pp.77–100). Cambridge, MA: MIT Press.

Sieck, W., & Yates, J.F. (1997). Exposition effects on decision making: Choice and confidence in choice. *Organizational Behavior and Human Decision Processes*, *70*, 207–219.

Sloman, S.A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22.

Slovic, P. (1995). The construction of preference. *American Psychologist*, *50*, 364–371.

Slovic, P., & Tversky, A. (1974). Who accepts Savage's axiom? *Behavioral Science*, *19*, 368–373.

Slugoski, B.R., Shields, H.A., & Dawson, K.A. (1993). Relation of conditional reasoning to heuristic processing. *Personality and Social Psychology Bulletin*, *19*, 158–166.

Slugoski, B.R., & Wilson, A.E. (in press). Contribution of conversation skills to the production of judgmental errors. *European Journal of Social Psychology*.

Smith, S.M., & Levin, I.P. (1996). Need for cognition and choice framing effects. *Journal of Behavioral Decision Making*, *9*, 283–290.

Stankov, L., & Crawford, J.D. (1996). Confidence judgments in studies of individual differences. *Personality and Individual Differences*, *21*, 971–986.

Stanovich, K.E. (in press). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Stanovich, K.E., & West, R.F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, *127*, 107–136.

Stein, E. (1996). *Without good reason: The rationality debate in philosophy and cognitive science*. Oxford: Oxford University Press.

Straughn, C.S., & Straughn, B.L. (Eds.) (1995). *Lovejoy's college guide* (23rd Edn.). New York: Simon & Schuster.

Takemura, K. (1992). Effect of decision time on framing of decision: A case of risky choice behavior. *Psychologia* , *35* 180–185.

Takemura, K. (1993). The effect of decision frame and decision justification on risky choice. *Japanese Psychological Research*, *35*, 36–40.

Takemura, K. (1994). Influence of elaboration on the framing of decision. *Journal of Psychology*, *128*, 33–39.

Thagard, P., & Nisbett, R.E. (1983). Rationality and charity. *Philosophy of Science*, *50*, 250–267.

Thaler, R.H. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, *1*, 39–60.

Tversky, A. (1975). A critique of expected utility theory: Descriptive and normative considerations. *Erkenntnis*, *9*, 163–173.

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

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458.

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

West, R.F, & Stanovich, K.E. (1997). The domain specificity and generality of overconfidence: Individual differences in performance estimation bias. *Psychonomic Bulletin & Review*, *4*, 387–392.

Willingham, W.W., Lewis, C., Morgan, R., & Ramist, L. (1990). *Predicting college grades: An analysis of institutional trends over two decades*. Princeton, NJ: Educational Testing Service.

Wolford, G., Taylor, H., & Beck, J.R. (1990). The conjunction fallacy? *Memory & Cognition*, *18*, 47–53.

Yates, J.F., Lee, J., & Shinotsuka, H. (1996). Beliefs about overconfidence, including its cross-national variation. *Organizational Behavior and Human Decision Processes*, *65*, 138–147.