ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES
Vol. 61, No. 3, March, pp. 305–326, 1995

Interaction with Others Increases Decision Confidence but Not Decision Quality: Evidence against Information Collection Views of Interactive Decision Making

CHIP HEATH AND RICH GONZALEZ

University of Chicago and University of Washington

We present three studies of *interactive decision making*, where decision makers interact with others before making a final decision alone. Because the theories of lay observers and social psychologists emphasize the role of information collection in interaction, we developed a series of tests of information collection. Two studies with sports collection show that interaction does not increase decision accuracy or meta-knowledge (calibration or resolution). The simplest test of information collection is responsiveness—that people should respond to information against their position by modifying their choices or at least lowering their confidence. Studies using traditional scenarios from the group polarization literature show little responsiveness, and even "deviants," who interact with others who unanimously disagree with their choice, frequently fail to respond to the information they collect. The most consistent finding is that interaction increases people's confidence in their decisions in both sports predictions and risky shift dilemmas. For predictions, confidence increases are not justified by increased accuracy. These results question theories of interaction which assume that people collect information during interaction (e.g., Persuasive Arguments Theory). They also question the labeling of previous results as "shifts" or "polarization." We suggest that interaction is better understood as *rationale construction* than as *information collection*—interaction forces people to explain their choices to others, and a variety of previous research in social psychology has shown that explanation generation leads to increased confidence. In Study 3, we provide a preliminary test of rationale construction by showing that people increase in confidence when they construct a case for their position individually, without interaction. © 1995 Academic Press, Inc.

We thank the following colleagues for their helpful comments and suggestions on earlier drafts: Max Bazerman, David Dunning, Bill Fulton, Linda Ginzel, Dale Griffin, Tony Greenwald, Beta Mannix, Terry Mitchell, Leigh Thompson, and Amos Tversky and two anonymous reviewers. Address reprint requests to Chip Heath, Graduate School of Business, The University of Chicago, 1101 East 58th Street, Chicago, IL 60637.

Decisions in the real world are often made *interactively;* we consult with others before deciding what jobs to take, what cars to buy, and what changes to make in our personal relationships. In such decisions, we collect information and opinions from others but make our final decision alone. We often follow this process when making personal decisions, but similar processes occur in other life domains. For example, because of norms about managerial autonomy in organizations, it is likely that most managers have traditionally followed an interactive procedure when making decisions—soliciting information from subordinates, peers, and superiors but making the final decision themselves.

People say that interaction plays a large role in important life decisions. When we asked a group of university students ($N = 118$) to think about important decisions that they faced (e.g., deciding what school to attend or changing a romantic relationship), almost everyone (91%) said that they sought the counsel of others while making such decisions. People also state that interaction helps them make more informed decisions. When the students were asked to rank order a list of reasons for interaction, the two responses that tied for first place concerned the collection of information. Subjects responded that (a) interaction gives them more information to make a better decision, and that (b) it causes them to think about things they had not considered (see Table 9 for other reasons and ranks).

The responses on this interaction survey are consistent with the conclusions of psychological research over the past 20 years in the literature on risky shift (Cartwright, 1971) and group polarization (Myers & Lamm, 1976). As we will discuss below, the most favored explanation of these phenomena (see e.g., Isenberg, 1986) has been that interaction allows people to collect information, and that this new information causes them to revise their decisions. "Shifts" or "polarization" are thought to occur "because during dis-

0749-5978/95 $6.00

cussion individuals are exposed to persuasive arguments providing information not available to them prior to making their initial choice" (Burnstein & Vinokur, 1975, p. 415).

The results of the interaction survey and the arguments of the social psychology literature suggest that people interact for the purpose of information collection. When we started our program of research we held a view not too different from the information collection view. However, several years of research in our labs have caused us to question this self-evident assumption about interactive decision making.

According to the studies in this paper, it appears that interaction is relatively ineffective as information collection. Decision quality does not improve much after interaction. Furthermore, people often resist modifying their decisions in response to the information they collect—even when confronted with the unanimous disagreement of their conversation partners, a situation when information collection views would predict they should be most likely to change. While providing little evidence of information collection, our studies do show that interaction produces robust and consistent increases in people's confidence in their decisions. Our results seem to have been anticipated by Ambrose Bierce's definition of conversation in his cynical *New American Dictionary* (1906). Bierce defined *conversation* as "a fair for the display of one's own monumental commodities; each visitor being too intent on the arrangement of his own wares to attend to those of his neighbor."

After presenting evidence against the information collection hypothesis, we attempt to understand why confidence may increase in the absence of increases in decision quality. We suggest that interaction may be better interpreted as an opportunity to "arrange our own wares" and develop a more coherent rationale for our decisions, than as an opportunity to collect information. We develop a view of interaction, the *rationale construction hypothesis,* that assumes that people increase their confidence after interaction because interaction prompts them to generate explanations for their choices and beliefs. This alternative view of interaction derives from a recent literature in social psychology (reviewed by Anderson & Wright, 1988; Koehler, 1991) that has demonstrated that people become more confident in their beliefs when they explain them. Interaction provides a natural context where people are prompted to explain their beliefs.

PREVIOUS RESEARCH ON INTERACTIVE DECISION MAKING

There has long been interest in both individual (Kahneman, Slovic, & Tversky, 1982; Nisbett & Ross, 1980) and group decision making (Hastie, 1986; Hill, 1982; Janis, 1972). In this paper we will concentrate on *interactive* decision making. We define interactive decision making as a procedure where individuals *consult with others but make their final decision alone.* Sniezek and Buckley (in press) have been interested in a similar issue in their studies of their Advisor–Judge model. Interaction gives people the opportunity to collect information and opinions from others, but because they make their final decision individually, they can use or ignore the information they collect during social interaction.

Why Study Interactive Rather than Group Decision Making?

In the past, two procedures have been used to study decisions in social environments. The first would be classified as "interactive" based on the definition above. The second is a true group procedure that requires groups to reach a consensual decision. Although it is reasonable that the two procedures produce different outcomes, the literature has treated them as interchangeable. In fact, most of the information we have on *interactive* decision making comes from research labeled as *group* polarization, *group* decision making, *group* judgment, etc. For example, Hastie (1986) summarizes a number of studies using both procedures that have been used in the literature on "group accuracy." In the group polarization and risky shift literature, the interactive procedure was used, for example, in Kogan and Wallach's (1967b) "*Group* Risk Taking" and in Burnstein and Vinokur is (1978) "Depolarization of Attitudes in *Groups*." In this paper, we use the label "interactive decision making" or "interaction" to avoid confusing interaction with the consensual group procedure.

We chose to concentrate on interactive and not group decision making for reasons of ecological validity and conceptual interest. In terms of ecological validity, the study of interactive decision making is important because many of our important decisions are made interactively. Although the majority of our decisions may be made individually, the interaction survey indicates that the majority of *important* decisions are made after significant social interaction. Interaction seems to be a better model of this social process than group decision making—consensual group decisions are comparatively infrequent in daily life. Conceptually, there are reasons to believe that interactive decision making may be very different from group decision making. In making group decisions people are explicitly aware of the trade-offs, negotiation, and logrolling necessary to reach a decision—only one decision will be made. This may change the dynamics of the social process—people may engage in relatively more problem-solving because they know that a common agreement must be found, or they may engage in more persuasion attempts because they know a common agreement must be forced.

As an important side benefit, the study of interactive decisions simplifies some complicated methodological issues associated with the study of group decisions. Groups must use some rule or procedure to aggregate individual opinions into a single group decision (e.g., perhaps an implicit or explicit majority rule vote). This aggregation procedure may hide or distort changes in individual preferences. For example, some researchers pointed out that when consensual group decisions are required, group decisions might "polarize" even if no individual changed his or her response (e.g., Abelson, 1973; Davis, 1973). The study of interactive decision making allows us to examine the impact of a social context on individual decision makers, while avoiding the complication of an unknown group decision procedure that may hide or obscure preferences.

Theories of the Effect of Social Interaction

Even though the group polarization and risky shift literatures have typically been labeled as "group decision making," these studies often used an interactive procedure. Since previous researchers did not distinguish between group and interactive decisions, we will assume that theories developed to explain "group" polarization were also assumed to apply to interactive decisions. This group polarization and risky shift literatures produced two distinct theories about the effects of interaction—one emphasizing information collection and the other emphasizing social comparison (see Isenberg, 1986, for a good summary of these theories).

Social Comparison Theory (Brown, 1965; Baron & Roper, 1976) assumes that people are motivated to view themselves in a socially favorable light. Individuals think well of themselves to the extent that they are distinct from the mass of humanity in the "right direction and to the right degree" (Brown, 1965). Group discussion provides information about how others feel, and after interaction, individuals revise their positions on the issues in order to keep their preferred position vis à vis the group. When all members of a group engage this process, the result is an overall shift in the valued direction. Social Comparison Theory in the risky shift literature differs from traditional social comparison research by emphasizing that people want to be "distinct" from their peers in the socially valued direction. Other research on social comparison also assumes that people use social information to confirm or verify their own responses (e.g., Fazio, 1978). All social comparison theories assume that people react to information that is generated by social interaction; however, they place their main emphasis on the social comparison process that drives reactions.

The more prominent theory explicitly invokes information collection. Persuasive Arguments Theory (Bishop & Myers, 1974; Burnstein, 1982; Myers & Bishop, 1971; Vinokur & Burnstein, 1978) assumes that an individual's position on any given issue will be a function of the number and persuasiveness of available arguments. People are assumed to evaluate the arguments favoring and opposing an alternative and to select the choice that collects the greatest number of arguments in its favor. These theorists assume that there is a cultural pool of available arguments that favor an alternative. Individuals come up with a few of these arguments on their own, but during interaction they collect novel arguments that favor the alternative they initially prefer, and they shift their initial opinions as the balance of arguments grows more extreme. Persuasive Arguments Theory stresses the importance of information collection: individuals should not shift unless they are confronted with new but valid arguments (Burnstein, 1982; Vinokur & Burnstein, 1978; Isenberg, 1986; Myers & Lamm, 1976).

In general, the literature has favored Persuasive Arguments Theory over Social Comparison Theory (Burnstein & Vinokur, 1973, 1975, 1978; Laughlin & Earley, 1982; Isenberg, 1986). Thus, given the responses of lay observers in the interaction survey, we seem to have a consensus between psychological theory and lay observation: Interaction is about information collection.

PREDICTIONS OF INFORMATION COLLECTION THEORIES

Because theories of interaction highlight information collection as the major goal of interaction, we developed below a series of tests, from relatively rigorous to least rigorous, to test whether information is collected during interaction.

Quality of Decision Test

The most rigorous possible test of the information collection hypothesis, because it involves the bottom line of decision-making, says that information collection during interaction should increase the quality of decisions made following interaction. This hypothesis is explicit in people's claims in the interaction survey—they claim that interaction allows them to collect more information and make better decisions. Persuasive Arguments Theory does not make specific predictions about quality of decisions because it was developed to explain reactions to moral and ethical dilemmas where it is hard to gauge "correctness." However, Persuasive

Arguments Theory does emphasize that it is important that arguments must be valid, relevant, and novel in order for them to influence decisions. Presumably, Persuasive Arguments theorists would agree that increasing the number of novel, relevant arguments would increase the quality of decisions.

Two of our studies examine predictions of future, real world events, a situation that should be advantageous for detecting effects of information on decision quality. If decision makers gain information during interaction, then this information should increase the accuracy of their predictions following interaction. Some studies have shown that interaction improves accuracy (e.g., Yates, 1991), but they have predominantly used general knowledge questions, a situation where truth, once stated, may be especially likely to win. The current studies will extend the test of decision quality to domains of predictions, where people must weigh knowledge in trying to predict a future that cannot be validated by previous knowledge.

Tests of Meta-knowledge

While increased accuracy would provide the simplest and most straightforward evidence of decision quality, the information available in the world, even if efficiently collected, may not always be sufficient to improve decisions. A slightly less rigorous test might only require that interaction allow people to more accurately assess the quality of the information they have. When we interact with others, we are presented with different viewpoints and are forced to consider critiques of our information and analysis (whether or not we would prefer to). Under this type of test, we would expect interaction to improve decision makers' *meta-knowledge* that is their judgments about the quality of their information.

Two tests that fall under this domain are tests of *calibration* and *resolution*. Calibration assesses how well subjects use the language of the probability scale; for example, when people assign events a probability of 70%, do those events happen about 70% of the time? Resolution assesses how consistently subjects sort their responses into groups—for example, is a game with a stated probability of 70% more likely to be won than a game with a stated probability of 65%? Calibration and resolution are independent: an individual can be calibrated but exhibit poor resolution, or exhibit perfect resolution but be imperfectly calibrated (Liberman & Tversky, 1993; Yates, 1990; Yaniv, Yates, & Smith, 1991). In the literature on individual decisions, there has been a long-standing interest in the topic of calibration (Fischhoff, Slovic, & Lichtenstein, 1977; Dunning, Griffin, Milojkovic, & Ross, 1990) and a relatively recent interest in resolution (Liberman & Tversky, 1993; Yaniv *et al.*, 1991). Resolution may provide the simplest test of meta-knowledge because it only requires people to distinguish between two groups of events—one that is more and one that is less likely to happen. Calibration adds the requirement that people's subjective probability match the relative frequency of events.

Evidence from studies of calibration has indicated that individuals are not very good at knowing what they know. People's expressed confidence in their judgments is typically greater than is warranted by their performance. Different researchers have speculated that this "overconfidence" is the result of people being (a) insufficiently critical of their own inference processes, (b) insufficiently aware of the fallibility of their memory and their ability to reconstruct information (Fischhoff, Slovic, & Lichtenstein, 1977), or (c) insufficiently attentive to information about the situation—for example, the base-rate of similar responses in the environment (Lichtenstein, Fischhoff, & Phillips, 1980; Dunning *et al.*, 1990). If interaction is an effective means of information collection, then it may correct these problems of insufficient information processing and as a result reduce overconfidence.

Some empirical evidence is not consistent with the view that interaction will increase meta-knowledge: Using general knowledge questions, Seaver (1977; cited in Von Winterfeldt & Edwards, 1986) found that individuals who participated in an interaction and a group consensus procedure were *more* overconfident in their predictions than individuals who participated in a no-interaction control condition. Seaver summarized his results by saying that "interaction among the assessors produces only a feeling of satisfaction and not any overall improvement in the quality of the assessed probabilities." Similarly, Dunning and Ross (1985) found more overconfidence following interaction using a prediction task about social behavior. Following the work of Dunning *et al.* (1990), these authors had students predict the behavior of a set of target students after they heard the targets respond to a series of questions that had been rated as being very useful in getting to know another person. For each target, subjects predicted a variety of actual behaviors (e.g., did the target read a book for pleasure last term?) and social responses (e.g., a good friend of the target is considering engagement to a person the target doesn't like, would the target say something?). Predictions made after interaction were significantly more confident and more overconfident than individual predictions. Although such increases in overconfidence are not universal (e.g., Hastie, 1986, cities two studies where interaction improves individual performance), there is some reason to doubt that interaction will automatically improve meta-knowledge.

Responsiveness Test

The least rigorous test of information collection says that information against one's position will reduce confidence in decisions. This test does not specify the magnitude of the reduction, merely that people's beliefs will be less strong after receiving information against them. This test is useful (although coarse) because it can be tested on the kind of qualitative decisions that have been used in the risky shift literature. According to Persuasive Arguments Theory, because individuals receive new arguments about their choices during interaction, we would expect that on average, choices would shift toward the majority alternative: during discussion, people holding a minority position receive on average relatively more arguments in favor of the majority alternative than people in the majority will receive in favor of the minority position. Although the responsiveness test is necessarily crude, if the information collection story is correct, these additional arguments should tip the balance for at least some people holding a minority position and should cause them to change their decision.

Present Studies

To explore the impact of information collection on post-interaction decisions, we present three studies that compare decisions made alone to decisions made following interaction with others. In keeping with previous literature, our experiments focus on interaction among equals. Our participants do not differ in expertise or status and do not in general have strong social ties with each other. While these factors are important in the real world, we wanted (like the previous literature) to assess the impact of information in situations where people share equal information and expertise and where they face no severe emotional consequences from either accepting or ignoring advice.

The informational role of interaction should be most important in situations where we can pool our knowledge with others' to understand or predict the external environment. In Studies 1 and 2, we test the effects of interaction on predictions about future real-world events (football games). This allows us to explore the accuracy of subjects' choices after interaction and to evaluate their meta-knowledge. In Study 3, we will use risky shift-style choice dilemmas to further examine information collection. Although we cannot test accuracy or meta-knowledge on choice dilemmas, we can test responsiveness—whether people are willing to change their decisions when interaction provides them with information against the option that they initially preferred.

STUDY 1: FOOTBALL PREDICTIONS

Study 1 explores the impact of interaction on decision accuracy and decision meta-knowledge. The information collection hypothesis predicts that both will improve as a result of interaction. The current study extends the existing literature on interaction by examining what happens when knowledgeable subjects interact to predict future events. Relatively few studies of interaction have assessed predictions of such events. This study also extends the literature on meta-knowledge by analyzing measures of resolution as well as calibration. Finally, we assess whether people are more willing to act (bet) on their predictions after interaction.

Method

Subjects were specially selected for their knowledge of the prediction domain. We advertised in an introductory psychology class for football fans to participate in a football prediction experiment and had them sign up in pairs. Twenty-one subjects participated in this study with their partners for 3 to 5 weeks in exchange for course credit in introductory psychology.

Each week, subjects individually predicted the outcome of 14 college and professional football games that were to be played the following weekend. For each game, subjects chose a winner and stated the probability that their chosen team would win. These measures will allow us to measure accuracy, calibration, and resolution in our analysis below.

Subjects also stated whether they would prefer to bet on their prediction about the game or a matched chance lottery. The matched lottery was a chance device that yielded a win with a probability equivalent to their confidence rating. Thus if a subject picked Dallas over Washington at 75% probability, the lottery was defined as drawing a number between 1 and 75 from a bag filled with 100 numbered poker chips. Lotteries were equated to confidence ratings at each stage of the game. If a subject increased his or her confidence by 5% after interaction, then the probability of success for the lottery increased by 5%.

After subjects turned in their individual predictions, they were given a separate form that contained seven games randomly selected from the 14 games they had already predicted. Subjects were told that the interaction period was designed to allow them to "share information and opinions" with their partner, but that they were not required to "reach a consensual decision" with their partner. Subjects interacted with their partner for as long as they wished (the typical pair interacted for a half an hour to 45 min), then they made new predictions about the outcome of these seven games.

Winning percentages were displayed weekly after the football games were played, and the participants with the best records earned (or seized) bragging rights among the participants.

Results

This procedure allows us to make within-subject comparisons on accuracy, calibration, and resolution, comparing individual responses before and after interaction. If information is being exchanged effectively during interaction, we would expect to see participants make more accurate predictions after interaction. In fact, predictions made after interaction were *not* more accurate than individual predictions ($M = +1.0\%$, n.s.). Although accuracy did not increase, *confidence* increased significantly. ($M = +2.9\%$, $t(2) = 2.89$, $p < .01$, pooling across weeks within subjects).[1]

To test meta-knowledge, we can examine calibration and resolution. Subjects were initially overconfident in their predictions. Although their confidence ratings indicated that they expected to be correct 66.2% of the time, their predictions were only 61.1% accurate, leading to a 5.1% discrepancy between expectations and performance. Because confidence increased more than accuracy, interaction exacerbated subjects' initial overconfidence. Figure 1 plots the isotone regression of the calibration curves for each condition. Overconfidence is signalled by the divergence of stated probabilities from actual probabilities (i.e., by the deviation of the calibration curve from the 45° line). As can be seen in this figure, the decrease in calibration following interaction occurs for all levels of probability, although it is especially pronounced for predictions made with high confidence.

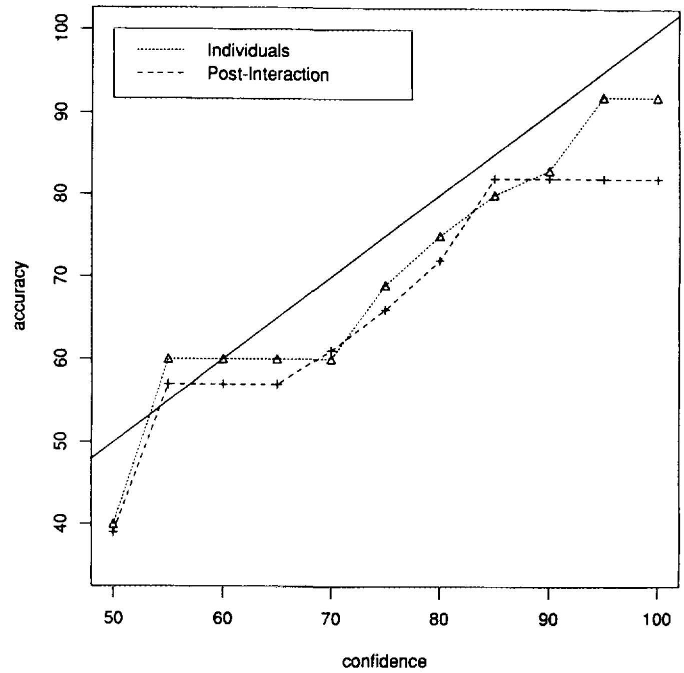

FIGURE 1

As another measure of meta-knowledge, we also calculate gamma as a measure of resolution. Gamma assesses how well subjects discriminated between games that were won and lost. It is defined as the percentage of concordances (stating one prediction with higher confidence than another, and having a corresponding difference in outcome) minus the percentage of discordances (having the prediction with lower confidence turn out to be true and not the prediction with higher confidence). Gamma can range between -1 (perfect negative association) and 1 (perfect positive association). The measure of resolution showed no change after interaction (Med $\gamma = .20$ both pre- and post-interaction).[2] Thus this index of discrimination suggests that subjects following interaction are not better at assigning higher confidence to those predictions that are more accurate.

The measures we report above provide useful tests of decision quality and meta-knowledge because they are directly interpretable within a calibration paradigm. However, they underrepresent the consistency of confidence increases after interaction. Therefore, we present two other analyses, one based on a subject-level analysis (whether each subject increased or decreased confidence after interaction) and one based on a prediction-level analyses (considering each prediction independently). The subject-level analysis pools over weeks for each subject. Seventy-six percent of subjects increased their confidence after interaction ($p < .01$ by sign test) and 62% of subjects increased in *overconfidence* after interaction ($p = .06$). The prediction-level analysis treats each prediction as independent. Table 1 displays changes in confidence versus changes in accuracy with each prediction contributing one entry. Although subjects gave different responses after interaction about 20% of the time, the responses were about as likely to move in the wrong direction as in the right direction. However, in contrast to this

[1] In our statistical analyses, we will typically try to report inferential statistics based on non-parametric tests, so that the significance levels of the test will hold for any monotonic transformation of the data. This kind of analysis will be especially important later in the paper when we perform tests on confidence measures for decisions in risky shift studies, since the confidence scales used in those studies do not have an immediate interpretation in terms of probability. Here, in keeping with the traditional calibration literature, we perform a test of overconfidence. The test of overconfidence implicitly assumes interval-scale data, otherwise the difference between confidence and accuracy would not have meaning. Thus, in this situation we report the t statistic.

[2] Whenever we report gamma as a measure of resolution, we also calculated another measure suggested by Liberman and Tversky (1993). In each case, the pattern of results is identical.

TABLE 1

Study 1: Change in Accuracy by Change in Confidence for Individual Predictions

Accuracy	Confidence			Total
	Down	Same	Up	
Up	10	11	33	54
Same	111	134	209	454
Down	21	8	22	51
Total	142	153	264	

Note. Cell entries represent the number of predictions for which the specified pattern was observed (k = 559).

equivocal effect of interaction on accuracy, interaction produced robust increases in confidence, with confidence increasing about twice as often as it decreased ($p < .0001$ by sign-test under the assumption of independence).

Subjects were more willing to bet on their predictions after interaction. Recall that for each game subjects stated whether they would prefer to bet on their prediction or on a matched lottery. Overall, subjects increased the proportion of prediction-bets following interaction ($M = +3.7\%$, $t(20) = 2.2$, $p = .05$). Note that this is a stringent test of willingness to act, since the lottery was equated to the prediction at both stages of the interaction.

Discussion

The introduction listed a number of reasons why interaction might improve decision quality: it presents us with new information and analyses, which should improve decisions, and it confronts us with others' critiques of our information and analysis, which should improve meta-knowledge. However, despite reasons to expect improvements, subjects are not more accurate after interaction, nor do they have more accurate meta-knowledge—resolution remains constant and calibration becomes poorer.

Although there is little evidence of information collection, there is evidence that interaction is a confidence builder. Subjects are twice as likely to increase rather than decrease their confidence after interaction. The confidence increases are *not* justified by increased accuracy; in fact, confidence increases lead to increased *over*confidence as demonstrated by the calibration curve of Fig. 1. Consistent with their increased confidence, subjects are also more willing to bet on their predictions after interaction. Since the probability of the chance lottery is equal to subjects' stated probability for their prediction, we might expect that subjects would be indifferent between betting on either the lottery or the prediction. This would be the prediction of Subjective Expected Utility theory. However, research has shown that people prefer to bet on their predictions when they feel competent and knowledgeable (Heath & Tversky, 1991). Thus the greater willingness of subjects to act on their predictions after interaction is consistent with greater feelings of competence after interaction.

The lack of evidence of information collection is even more striking when we consider the prediction of the responsiveness test that confidence should decrease when partners disagree before interaction. Presumably, partners disagreed on their pre-interaction predictions because they had very different information in favor of the different teams—different enough to cause them to wind up with different preferred choices. Upon exchanging information, both partners should have new arguments to support the team they did not choose, and this additional information should lead them to moderate their confidence.

Even when partners *disagreed* on their pre-interaction predictions (Table 2), their confidence still tended to increase following interaction. In 88% of the cases where partners disagreed prior to interaction, both partners chose the same team after interaction. However, instead of moderating their confidence based on the informational differences signalled by the disagreement, *both* individuals tended to increase in confidence. Thus there tended to be one partner who adopted a new team with even more confidence than was assigned to the original team. We thus see responsiveness at one level (choices tend to respond to the information collected and partners tend to agree after interaction), but not on another (confidence does not respond to the information collected during interaction).

It might be possible to explain the confidence increases as a methodological artifact. Since subjects respond twice to the same materials it is possible that confidence increases are produced by familiarity (consider, for example, Tesser's (1978) work on 'mere

TABLE 2

Study 1: Change in Accuracy by Change in Confidence for Partners Who Initially Disagree

Accuracy	Confidence			Total
	Down	Same	Up	
Up	10	10	33	53
Same	37	38	35	110
Down	19	8	20	47
Total	66	56	88	

Note. Cell entries represent the number of predictions for which the specified pattern was observed (k = 210).

thought'). However, the previous interaction literature has argued convincingly that the effects of interaction cannot be explained by familiarity. Having people engage in a writing task does not produce the same effects as interaction (Bell & Jamieson, 1970; Teger, Pruitt, St. Jean, & Haaland, 1970), neither does responding a second time to materials after intervening thought about a separate topic (Burnstein & Vinokur, 1975). In this paper, Study 3 will use a between-subjects design and still show a similar pattern of confidence increases in a design that eliminates familiarity effects as an alternative explanation.

STUDY 2: INFORMATION COLLECTION INDIVIDUALLY AND DURING INTERACTION

Interaction in the first study did not improve decision accuracy or meta-knowledge. At the same time, it did produce robust confidence increases. Because these results are surprising for information collection theories, Study 2 was designed to extend and elaborate the results of the first study.

We again ask people to make predictions about football games and give their confidence in their prediction. To elaborate the results of the first study, we change our procedure to bias our test more in favor of information collection. In Study 2 we use non-expert subjects and increase the size of the interaction groups. Subjects who are less-informed about football should be more responsive to the information that is available in the environment, and subjects who interact with more people should have more of an opportunity to collect and pool information. The larger groups will also allow us to do an expanded test of the disagreement results of the previous experiment: how do decisions respond when an individual receives information from multiple people who disagree with him or her?

To extend the results of the first study, we add a condition which will allow us to explore what happens when we provide people with a very rich source of information—preseason team reports compiled by professional sports analysts. The reading condition is not intended to provide similar information to the interaction condition; instead, it is designed to allow us to contrast the effects of "information collection" during interaction with the effects of information collection from an extremely good individual source of information. For example, if confidence increases in the first study are driven by perceived information collection, then the reading condition should show much more striking increases in confidence because it allows people to collect more information.

Method

Subjects participated in the experiment in exchange for course credit in introductory psychology. The subjects were taken from the general subject pool and were not asked about their football expertise before signing up for the experiment.

During the experiment, subjects predicted the winners for 28 football games (14 professional and 14 college games) played on a given weekend. Subjects also rated their confidence in their prediction and their knowledge of the 28 games. We will refer to this part of the experiment as the "baseline phase."

After completing their responses, subjects were given a second list that contained 14 of the original 28 games. One of two experimental manipulations occurred at this point:

Interaction condition ($N = 41$). In this condition subjects interacted in groups of three to five participants, then responded again to the experimental measures. The interaction instructions were the same as in the previous study—subjects were told that the interaction period was designed to allow them to "exchange information and opinions with the other participants," but that they were not required to reach a common decision. Subjects generally started their discussions by going around the table and having everyone state their position along with the major reasons that they decided on their position. Disagreement in initial responses prompted more detailed discussion and elaborate explanations. Discussions typically lasted from 3 to 5 min for each game, with longer discussions prompted by disagreement. The experimenter acted as a roving moderator, singling out someone to start a discussion in the occasional quiet group, and urging people not to get bogged down in groups that had discussions that lasted a very long time. After interaction, subjects predicted the results of the 14 games on the second list a second time.

Reading condition ($N = 37$). Here subjects read written descriptions of the two teams prior to making their second set of predictions. The description of each team was photocopied from *Gameplan's Football Forecast*, a magazine published prior to the start of the football season. The printed material included the previous season's record, descriptions of returning starters, a written overview of both offense and defense, and predictions for season performance. Although the information was extensive, it did not involve predictions about specific games. After contemplation, subjects predicted the results of the 14 games on the second list a second time.

To provide an internal control, after participating in one of the two manipulations, subjects predicted the results of the remaining 14 games a second time. We will refer to this final round of predictions as the "return-to-baseline" phase. This experiment is an ABA design: baseline, manipulation, and return-to-baseline, with each subject contributing 56 observations (28 games predicted twice). The two lists of 14 teams were counterbalanced across the manipulation and return-to-baseline phases.

Betting measure. Subjects were given the opportunity to bet on one of two predictions, one randomly chosen from the manipulation phase and one from the return-to-baseline phase. Subjects won $1 if the team they selected from the pair of games actually won the game.

Rose Bowl. At the end of the experiment, we also collected information on a game which held special meaning to our University of Washington subjects. Subjects were asked to predict the winner of the 1991 Rose Bowl, which pitted Washington against Iowa. This study was run during the week when the Washington qualified for the Rose Bowl, but their opponent had not yet been named; thus, subjects tended to have very little information about Iowa. Subjects were told that Washington had qualified for the Rose Bowl and were asked to imagine that the Iowa would be their opponent. Fortunately, the real world cooperated and Iowa earned the Rose Bowl spot the following week.

Since the opportunity to collect data on this game occurred after our other experimental materials had been prepared, we ran the procedures for the Rose Bowl in a between-subjects design. Two groups of subjects chose a team and rated their confidence after either an interaction ($N = 28$) or a reading manipulation ($N = 28$). An additional group of subjects ($N = 22$) served as a control and simply responded to the experimental measures (i.e., they did not participate in a group discussion nor did they read information).

Results

In our analysis of the current study, we will report two kinds of results: As in Study 1, the overall results for the sample will again allow us to assess accuracy and meta-knowledge. In this study, we will be able to supplement these results by examining the same measures for the reading condition. Second, we expand the analysis of disagreement conducted in Study 1 by examining the behavior of individuals in situations where they are a lone deviant interacting with people who unanimously disagree with them. The clear prediction of information collection is that deviants will modify their opinions after being exposed to information and arguments against their position.

First we consider tests of accuracy and meta-knowledge. In contrast to the last experiment, subjects in this experiment were not specifically selected for their football expertise and the sample of games predicted was much larger. Although their mean confidence is similar to the more expert subjects of Study 1 (67.8% for baseline predictions), their mean accuracy (49%) is much lower than in the previous study. Accuracy increases in the reading condition (Wilcoxon signed-rank test, $p < .0001$, with each subject's percentage correct as the dependent variable) but not in the interaction condition ($p = .15$). There are no significant differences between the conditions in the return-to-baseline phase. As in Study 1, the predictions of information collection theories are not confirmed on the most rigorous test—interaction does not increase accuracy.

Moving to tests of meta-knowledge, we observe that baseline predictions are tremendously overconfident (a difference of 18.8% between expected frequency of success and actual success). Both reading and interaction subjects increase in confidence during the manipulation phase (Wilcoxon signed-rank test, $p < 0.001$). However, overconfidence decreases significantly for reading subjects, while it remains essentially the same for interaction subjects (the insignificant increase in accuracy for the interaction subjects cancels out the increased confidence). During the manipulation phase, average confidence for interaction subjects is 71% and their accuracy is 53% (a difference of 18%), yielding similar overconfidence to baseline. Average confidence for subjects in the reading condition is 72% and their accuracy is 60% (a difference of 12%). Figure 2 provides the calibration plots for the two conditions before and after the experimental manipulation.

Table 4 records resolution (gamma) for each condition during each phase of the experiment. Compared to interaction, the reading manipulation is more successful in allowing subjects to discriminate between games that will be won and games that will be lost (Mann–Whitney $p < .03$ for manipulation phase, all other differences n.s.).

Table 3 records for this experiment the prediction-level analysis of confidence and accuracy performed in the previous study. Again, confidence is about twice as

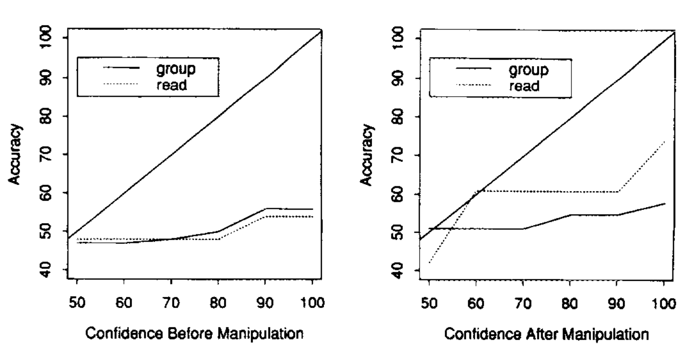

FIGURE 2

likely to increase as to decrease ($p < .0001$ by sign-test under the assumption of independence).

Betting measure. We used the betting measure to examine whether people are more willing to act on their beliefs after interaction. Recall that subjects were given the opportunity to bet on their prediction for a game from the manipulation phase or a game from the return-to-baseline phase. Subjects received $1 if their favored team won the game they chose. Unfortunately, the random procedure used to select the games produced one game which involved a team in our subjects' home conference, and therefore was much more familiar to our subjects than the other game. Subjects preferred to bet on the game with the more familiar team, but the preference was influenced by the phase in which the game appeared: 94% chose to bet on the game when it appeared during the manipulation phase, but only 71% preferred the same game when it appeared during the return-to baseline phase ($z = 5.9$, $p < .001$). There were no differences between reading and interaction conditions, and both manipulations increased subjects' willingness to bet on their predictions.

Rose Bowl. We added the Rose Bowl game to our materials when the University of Washington qualified for the Rose Bowl during the week of the experiment, and this game was run in a between-subjects design with an interaction condition, a reading condition, and a control group. An independent group of Washington undergraduates rated the reading materials as being about equally favorable for each team. Not surprisingly, almost all participants (96%) favored

TABLE 3

Study 2: Change in Accuracy by Change in Confidence for Interaction and Reading Conditions

	Confidence			
Accuracy	Down	Same	Up	Total
Interaction condition ($k = 574$)				
Up	18	32	34	84
Same	111	137	189	437
Down	13	15	25	53
Total	142	184	248	
Reading condition ($k = 517$)				
Up	23	23	56	102
Same	95	105	173	373
Down	9	12	21	42
Total	127	140	250	

Note. Cell entries represent the number of predictions for which the specified pattern was observed.

TABLE 4

Study 2: Comparing Gamma Measure of Resolution for Interaction versus Reading Conditions

	Baseline	Manipulation	Return-to-baseline
Interaction	.05	.03	.02
Reading	−.03	.17	.08
	n.s.	$p = .03$	n.s.

Note. The number reported in the median gamma for the specified phase of the experiment.

Washington in their initial predictions.[3] After the experimental manipulation, subjects in the interaction condition expressed more confidence in their chosen team ($M = 85$) than subjects in the reading condition ($M = 77$). The control subjects fall in between ($M = 80$). The only pairwise comparison to achieve statistical significance is interaction vs reading (Mann–Whitney U, $p = .02$).[4]

Analysis of Disagreement

We now present an internal analysis of the effects of interaction when there is a deviant (a minority of one). Note that this provides the strongest test of information collection theories: based on information collection we should expect everyone to switch after being exposed to the arguments of several others who disagree. To create a comparison group for the interaction condition, we analyze the responses of individuals who participated in the reading condition; these subjects participated concurrently but did not interact.

Table 5 displays changes in confidence versus changes in accuracy for deviants in both interaction and reading conditions. Both manipulations caused people to change their answers. Fifty-seven percent did so in the interaction condition versus 43% in the reading condition. The difference between interaction and reading conditions is nonsignificant.

It is not surprising that an individual who is confronted with the unanimous opposition of others would change his or her answers. What *is* somewhat surprising is the number of deviants who held on to their answers. Twenty-four percent (24.4%) of deviants held to their original response with the same or increased confidence.[5] It is difficult to know how to state the null hypothesis in this situation, so we simply report a con-

[3] Washington in fact won the game, leading to high (but perhaps accidental) accuracy for this game. We regret we could not run the identical study at Iowa.

[4] Whenever we report between-subjects tests for interaction subjects, we also ran parallel analyses which account for nonindependence within the interaction conditions (see Myers, DiCecco, & Lorch, 1981). The results we report are unchanged by these analyses.

[5] No subject contributed more than one observation to this analysis.

TABLE 5

Study 2: Deviants Only—Change in Accuracy by Change in Confidence for Interaction and Reading Conditions

Accuracy	Confidence			Total
	Down	Same	Up	
	Interaction condition (k = 49)			
Up	3	6	8	17
Same	9	11	1	21
Down	0	4	7	11
Total	12	21	16	
	Reading condition (k = 44)			
Up	1	1	8	10
Same	4	9	12	25
Down	1	3	5	9
Total	6	13	25	

Note. Cell entries represent the number of predictions for which the specified pattern was observed.

fidence interval.[6] However, relative to expectations about possible sources of noise, the 95% confidence interval (15%, 39%) seems resolutely positive. Based on Persuasive Arguments Theory it is certainly surprising that one of four deviants fails to change even their confidence in their answer upon being confronted with unanimous disagreement.

Discussion

The current study again shows little support for information collection theories of interaction. The reading condition demonstrates that it is possible for the collection of information to improve decisions on this task. Note that the information in the reading condition raised performance of reading subjects almost to the levels of our expert subjects in Study 1 (as measured by both accuracy and resolution). Interaction, on the other hand, did not produce better decisions.

As in Study 1, we find that interaction produces robust increases in confidence. The ratio of confidence increases to decreases is almost two to one. If confidence increases were primarily driven by actual or perceived information collection, then subjects in the reading condition should have shown much greater increases than subjects in the interaction condition. Despite the fact that the two conditions involved very different real (and probably perceived) amounts of information collection, absolute confidence increases were essentially the same in the two conditions.

The results of the deviant test are quite surprising, and provide compelling evidence against the information collection hypothesis. It might be possible to salvage the information collection theories, for example, by arguing that the deviants might not have collected sufficiently novel, relevent, and valid information to produce any changes in their position. While this may be possible, it leads to an untenable position for information collection theories. Persuasive Arguments Theory, for example, assumes that people develop their *initial* opinions by sampling from a pool of available arguments, weighing arguments for and against a particular alternative, and deciding accordingly (Burnstein & Vinokur, 1975, 1977; Kaplan, 1977). Under this model, a deviant who confronts three other people is facing a large pool of persuasive and relevant arguments against his or her position, because trivial or irrelevant arguments would not have encouraged the majority to adopt their preferred alternative in the first place. After interaction with this majority, a deviant could refuse to modify his or her position only if *every* argument of the majority had already been considered (e.g., Kaplan, 1977; Vinokur & Burnstein, 1978). Thus, in making their initial decision, deviants not only have to have generated every argument generated by the majority for their position, but also a larger number of arguments on the opposite side of the issue. This implies that the deviant has much greater knowledge than the combination of his or her opponents, an assumption that is unlikely because subjects shared approximately the same level of expertise and because intransingent deviants showed no signs of being especially accurate.

INTERPRETING INCREASES IN CONFIDENCE

Because the information collection hypothesis was partially derived from the group polarization literature, we felt it was important to examine information collection theories in this domain. We particularly wanted to confirm two aspects of the current results: first, the relative insensitivity of decisions to information collection and second the tendency for confidence to increase after interaction. Although it is harder to specify clear tests of information collection in risky shift problems, the responsiveness test predicts that some people should shift their responses after interaction and this should be particularly true for deviants.

Studying risky shift dilemmas in addition to the prediction tasks of Study 1 and 2 will provide important information about whether the phenomenon we document is domain-specific. The two tasks model two very important classes of decisions that we face—football

[6] The null hypothesis for this percentage would probably be higher than zero, for example, because it would need to account for the probability of subjects making errors in recording their responses and experimenters making errors in transcribing responses. Thus, it wouldn't be appropriate to report a simple t-test for comparing the observed percentage to zero (the theoretical prediction of information collection theories).

predictions model situations of expertise, where there are many facts and aggregating information is important; risky shift dilemmas model the assessment of value in ambiguous social arenas, where arguments have less to do with assessing what the world will do, and more to do with assessing what we think is prudent or proper or moral.

In the next section, we first present briefly a study from a separate program of research that provides an initial test of responsiveness. This study fails to show responsiveness, but replicates the confidence increases of Studies 1 and 2. In response to the data of this study, we will take *post-interaction confidence increases* as the phenomenon to be explained and will consider a number of hypotheses to account for confidence increases. Then we will introduce the final experiment, which will serve both as a final test of information collection and a preliminary test of another possible model of interaction.

Preliminary Evidence: Changes in Confidence and Choice in Risky Shift Problems

Preliminary evidence about the impact of interaction on risky shift problems can be taken from a study from another research program. We will label this as the Baseline Shift study, and we will use the results from this study to motivate the discussion that follows.

This study changed the dependent measure used in the original risky shift studies to be more consistent with the binary measures used in the football prediction studies. In the original risky shift studies, subjects specified "the lowest probability of success" that they would consider acceptable in order to recommend that a person pursue a certain risky option. For example, in one problem, subjects were asked to recommend whether an engineer should leave a secure but dead-end job for a risky but higher potential job with a start-up firm. People used a 0–100% risk scale to state how likely the startup firm had to be to succeed in order for them to recommend a switch to the new firm. Responses could vary from 0% (i.e., the engineer should take the job even if there were no chance the new firm would succeed) to 100% (the engineer should take the job only if the new firm would certainly succeed). In the current study, instead of using the risk scale, we ask people to choose one of the two options implicit in each problem and then rate their confidence in their choice on a "0" (no confidence) to "10" (high confidence) scale.

In a between-subjects procedure, 118 subjects examined eight choice dilemmas: six from the original risky shift studies and two new ones (See Appendix for dilemmas). Subjects considered half of the choice dilemma in an individual condition and half in an interaction condition. The interaction condition again involved a free discussion among three to five participants with instructions and procedure similar to those of Study 2.

TABLE 6

Baseline Shift Study: Compare Shifts on Construals, Answers, and Confidence for Interaction Condition Subjects versus Individual Subjects

	Change construals	Change answers	Confidence
Surgeon[a]	−.06	−.07	.53*
Vacationer[a]	−.01	−.10	.25*
Football	.18	.00	.02
Company location	.07	−.07	.57*
Med School	−.20	−.14	.07
Engineer	−.14	.02	.34*
Student	−.11	.16	.07
Career	.09	−.14	.15
Cochran's z	.09	.04	6.6
Overall "p"	n.s.	n.s.	$\leqslant .001$

Note. For Change Construal and Confidence columns the cell entries are the differences in the proportion of responses above the median for individual versus interaction subjects. Positive values indicate that a higher proportion of interaction subjects are above the overall median. Entries in the Change Answers column are differences in the proportion of people choosing the "risky" alternative.

[a] These dilemmas have traditionally shifted conservative.

* Individually significant at $p < .05$.

Table 6 presents the between-subjects analysis of the choice and confidence measures. The responsiveness test would predict that some people would change their choices after interaction. Instead, participants made essentially the same decisions after interaction as they made prior to interaction. The aggregated results were nonsignificant (Cochran's $z = .04$; n.s.)[7] and no individual problem showed a shift. Although the choice measures give no indication that people's choices responded to the information they collect, participants show significantly more confidence in their decisions after interaction (Cochran's $z = 6.6$, $p < .001$). Thus, subjects appear to give the same answers after interaction but with much more confidence.

Rethinking the Effects of Interaction: The Rationale Construction Hypothesis

Although the studies thus far have provided little evidence that interaction involves information collection, interaction does seem to produce robust and consistent confidence increases. A similar pattern of results is found across two different domains (football predictions versus risky shift dilemmas) and using different response scales (probability ratings for football

[7] We combined individual χ^2s according to Cochran's method (1954) which converts individual χ^2s to z-scores and then aggregates signed z's. An identical pattern of significance is found for aggregation of individual χ^2s which does not take sign into account.

predictions and confidence ratings in the Baseline Shift Study). Additionally, these findings are consistent with previous findings that interaction exacerbated overconfidence (Seaver, 1977; Dunning & Ross, 1985). Thus post-interaction confidence increases seem to be the key phenomenon to be explained. In this section we consider different explanations for the confidence increases. We first explore two hypotheses that suggest that the social environment of interaction leads to increased confidence either by giving people the perception that their information has increased or by providing social agreement that validates their decisions. After pointing out deficiencies in these hypotheses, which locate the source of confidence in the external environment, we propose a hypothesis that locates the source of confidence internally, arguing that confidence increases result from a particular information-processing style prompted by the social environment. This discussion will then generate a particular hypothesis, the *rationale construction hypothesis,* which will be tested in Study 3.

The results of accuracy and meta-knowledge tests indicate that little information is collected during interaction. Despite this empirical result, however, one might argue that people believe that they collect information during interaction and increase their confidence based on this belief. We can label this the *perceived information gain* hypothesis. A vivid counterexample to this hypothesis is the response pattern in the Rose Bowl game in Study 2. Discussions for the Rose Bowl explicitly highlighted the fact that the information available to our University of Washington subjects was highly redundant (and biased for emotional reasons). When asked to predict the winner of this game, the interaction typically started with someone calling out [Washington] "Huskies" in a cheerleader tone and the other participants chiming in their agreement. Iowa was rarely mentioned, except occasionally when a participant would state that they knew nothing about Iowa (others would nod in agreement) or would ask if anyone knew about Iowa's record (no one did). The Rose Bowl discussions are thus a prototypical case where interaction *should not* increase confidence under the perceived information gain hypothesis—the information exchanged was highly redundant, and the interactions explicitly highlighted the fact that no one had any real information about Iowa.

The perceived information gain hypothesis assumes that people increase their confidence after interaction because they believe they have more *valid* information about the external environment. Alternatively, we might propose that people believe they have more *reliable* information after interaction. The *inference certification hypothesis* states that subjects increase their confidence, not because they have received novel information, but because they are more certain that they have correctly evaluated their existing information (cf., Dunning & Ross, 1985). When others agree with us, they confirm that we have made appropriate inferences based on the available information. Under this hypothesis, interaction does not provide us with novel information or arguments, but with reliability checks on our inferences. Similar predictions for situations of agreement are made by research on "validation" (Fazio, 1979) and by Social Comparison Theory (e.g., Meyers & Lamm, 1976). The predictions of the later theories derive from social forces more than from information collection, however, each of them requires a decision maker to collect some (limited) information about the responses of other people.

While inference certification might explain confidence increases after redundant information exchange, even this hypothesis would predict that deviants would lower their confidence after interaction. Instead, a notable fraction of deviants do not change their answers or lower their confidence after interacting with two to four others who unanimously disagreed with them. Thus both perceived information gain and inference certification hypotheses fail to account for the complete pattern of results in these experiments.

How then might confidence increases be explained? If interaction is not an opportunity to collect information from the external environment—either in the form of persuasive arguments (perceived information gain) or in the form of social information about other's positions (inference certification)—then perhaps confidence increases result from an internal process that is prompted by the social environment, but that does not require or respond to information from the social environment. This conception matches that of Ambrose Bierce, who claims that interaction is "a fair for the display of (our own) commodities" and that we are too busy during interaction "attending to our own wares to attend to those of our neighbor." It also matches our observations of the interaction process: when people in our experiments confront disagreement, they do not appear to spend much time trying to elicit and understand the arguments offered by others; instead, they appear to spend the majority of their time trying to patiently (or impatiently) explain their own view to others.

A large literature in social psychology has demonstrated that people become more confident in their beliefs when they explain them (e.g., Anderson, 1983; Campbell & Fairey, 1985; Hoch, 1985; Ross & Anderson, 1982; Sherman, Skov Hervitz, & Stock, 1981; Sherman, Zehner, Johnson, & Hirt, 1983; see Anderson & Wright, 1988, or Koehler, 1991 for reviews). These studies ask subjects, for example, to explain a relationship between causal variables or to explain why an event might happen. After generating an explanation, subjects become more confident that the relationship does exist or that the event will happen.

We argue that there is a logical parallel between the thought processes required by the experiments above and the thought processes required by social interaction. In the experiments above, an experimenter asks subjects to explain a particular belief, and after constructing an explanation for their belief, people are more confident in their belief. In our studies, interaction forces people to explain their decision to others, and after constructing an explanation for their decision, people are more confident in their decision. Because interaction provides a natural analogue of these explanation generation tasks, it is not surprising that our results on interaction parallel the results of the studies above.

Koehler (1991) summarizes the literature on explanation generation by arguing that confidence is a function of the fit and coherence of the case that people construct for their beliefs. A similar role of coherence is proposed by Pennington and Hastie (1992, 1988), who argue that jurors' confidence in their verdict will be determined by the coverage, coherence, and uniqueness of the story that they construct about the events discussed during the trial. Following the lead of these authors, we propose that as people organize their thoughts and articulate a coherent rationale for their choices during interaction, the increased organization and coherence of their case may make them more confident in their decisions. If this analysis is correct, we might better describe interaction as a process of *rationale construction* rather than as information collection.

Consistent with the rationale construction hypothesis are findings from the group polarization literature that passive receipt of arguments outside interaction produces smaller shifts than interaction (Bishop & Myers, 1974; Burnstein & Vinokur, 1973) and that listening to a discussion generates less shift than actual participation (Bell & Jamieson, 1970; Lamm, 1967). Active construction of a case seems to be necessary, and not just passive information collection. If the "rationale construction" view is accurate, then individuals who construct a rationale for their choices should show confidence increases similar to those of interaction subjects. Study 3 will test this assertion.

STUDY 3: INTERACTION AND ARGUMENTS

The current study thus has two purposes: to extend the test of information collection in the domain of risky shift style dilemmas and to provide a preliminary test of the rationale construction explanation for the confidence increases in our previous studies.

To extend the test of information collection, Study 3 examines the effects of interaction using the choice dilemmas of the original risky shift studies. In contrast to the predictions of responsiveness made by information collection theories, the Baseline Shift study indicates a surprising amount of stability in choices on risky "shift" dilemmas. However, the between subjects design of that study may be too insensitive to detect changes. The current study allows a more sensitive test by providing within-subjects as well as between-subjects comparisons. As a special test case of the responsiveness test, this study will again examine the impact of interaction on the choices of deviants using the analysis that we performed in Study 2. Information collection theories make a clear prediction that deviants should modify their responses after collecting arguments and information from the majority who disagrees—if deviants do not actually change their choice, they should at least make their choice with less confidence.

To provide a preliminary test of rationale construction, we attempt to mimic the effects of interaction using a pure information processing manipulation that does not involve interaction. Studies in the risky shift literature have often used such manipulations to explore whether information processing effects (e.g., familiarity) could account for the effects of social interaction (Bell & Jamieson, 1970; Teger *et al.*, 1970; Burnstein & Vinokur, 1975). In Study 3 we use an individual writing procedure that asks people to construct a case for their choice (the Directed Writing procedure). If rationale construction is sufficient to increase people's confidence in their choices, then this procedure ought to produce similar results to interaction.

To control for processing time and effects such as mere thought (Tesser, 1978), we ran another writing condition that provides an experimental control, but that also has a theoretical interpretation. Information collection theories in social psychology typically assume that people weigh arguments *independently* according to their relevance and validity, and choose the option that collects the greatest weight of arguments (e.g., Myers and Lamm, 1976). This process differs in two key ways from rationale construction. First, it assumes that people attend to arguments for *each* alternative; rationale construction assumes that people construct a case for the *preferred* alternative. Second, this process assumes that people consider and weigh arguments independently; rationale construction assumes that people assess the fit and coherence of the rationale they construct.

To test these differences in processing, we created another individual writing procedure which asks people to independently consider and weigh arguments. This procedure asks subjects to consider both the pros and cons of the available alternatives (the Pro/Con Writing procedure). Thus, the Pro/Con procedure, in addition to providing an experimental control for processing time and familiarity, will also allow us to examine whether the effects of interaction are more successfully modeled when people consider relevant arguments independently or when they construct a case for their decision. Variants of the Pro/Con procedure have been used in the past as controls, and typically they did not parallel the effects of interaction (Bell & Jamieson, 1970; Teger *et al.*, 1970; Burnstein & Vinokur, 1975). These researchers interpreted these results to indicate that the effects of interaction were produced by either information or social support from the external environment. If the rationale construction hypothesis is true, a better interpretation may be that the effects of interaction are *not produced* by the independent weighing of arguments.

The *information collection* hypothesis predicts the following ordering for the three conditions on the responsiveness test: Interaction >> Pro/Con = Directed Writing. Interaction ought to produce much larger effects than either writing procedure because writing subjects do not collect information.

Rationale construction predicts that decisions will be relatively unresponsive to information exchanged in interaction, and thus that there will be no systematic differences in responsiveness between conditions: Interaction = Pro/Con = Directed Writing. However, both interaction and Directed Writing should produce confidence increases because they encourage rationale construction. The Pro/Con procedure should not increase confidence because it prevents rationale construction by encouraging people to attend to arguments against their choice as well as for their choice. Thus the rationale construction hypothesis predicts that for confidence: Interaction = Directed Writing >> Pro/Con.

Method

Subjects were 255 undergraduates participating for course credit or extra credit in introductory psychology courses. Subjects examined six dilemmas from the Choice Dilemmas Questionnaire used in risky shift studies (Kogan & Wallach, 1967). Of the six dilemmas we selected from the original group of 12, two have historically exhibited cautious shifts (Vacationer, Surgeon) and four have typically exhibited risky shifts (Career, Chess Player, Engineer, and Company Location). For each of the six dilemmas, subjects chose one of the two courses of actions, rated their confidence in their response, and then responded to the traditional risky shift question using the traditional scale.

Subjects responded to three dilemmas in an interaction condition and three in a writing condition. The order of the interaction and writing tasks are counterbalanced across conditions, and the experimenter ensured that subjects spent approximately the same amount of time on both tasks. The interaction condition was defined in the same way as in previous studies. The writing condition differed across groups:

Pro/con writing ($N = 142$). In this manipulation, subjects generated arguments for and against each alternative action. Each choice dilemma was presented on a separate page, and below the vignette, the page contained four separate spaces in which subjects were to write arguments for and against Option A and B of the choice dilemma (spaces were labeled "Arguments for A," "Arguments against A," etc.).

Directed writing ($N = 113$). In this manipulation, subjects were asked to construct a case in favor of their preferred alternative. Each choice dilemma was presented on a separate page and subjects were given space to generate and organize their case in support of the option they favored.

Approximately half of the subjects participated in a between-subjects procedure, recording their choices only once after the interaction or writing task. The remaining subjects participated in the traditional within-subjects procedure, recording their choices both before and after the experimental task.

Results

We will first examine the between-subjects comparisons and then turn to the results of the more sensitive within-subjects comparisons.

Between-subjects results. Table 7 shows the results of between-subjects tests, comparing initial individual responses from the within-subjects conditions with post-manipulation responses of subjects in each of the three conditions.

Consistent with the results of the Baseline Shift study, interaction did not produce changed answers, but did produce increased confidence ($p < .001$ by Cochran's z). The Directed Writing condition shows a similar pattern ($p < .001$ by Cochran's z). In contrast, the Pro/Con Writing procedure shows no shifts on any measure. No condition produced significant changes in risk. This failure to replicate the traditional risky shift will be discussed further in the discussion section.

Within-subjects results. Within-subject comparisons (see Table 8) confirm that subjects tend not to change their choices after interaction. Only 6.8% of decisions changed after interaction, 3.9% after the Pro/Con Writing procedure and .5% after Directed Writing. Analyzing confidence across problems within each subject, average confidence increases after the interaction

TABLE 7

Study 3: Between Subjects Analysis—Compare Shifts on Choice, Confidence, and Risk for Subjects in Three Experimental Conditions versus Baseline

	Choice			Confidence			Risk		
	Int	DW	P/C	Int	DW	P/C	Int	DW	P/C
Engineer	.08	−.12	−.04	.25*	.30*	−.04	.19*	.04	−.16
Chess player	−.04	−.04	.08	.18*	.02	−.02	−.04	−.09	−.16
Company location	−.10	−.09	.03	.35*	.14	−.05	−.11	−.14	−.02
Career Decision	−.08	.06	−.01	.13	.08	.05	−.05	.00	−.05
Surgeon[a]	.04	.14	.02	.17	.19	.12	.11	.01	−.01
Vacationer[a]	−.11	−.22*	.09	.25*	.36*	−.06	.14	.03	.03
Mean	*−.04*	*−.05*	*.03*	*.22*	*.18*	*.00*	*.04*	*−.03*	*−.06*
Cochran's z	−.29	−1.5	.90	6.3	3.6	.56	−1.1	.58	1.4
Overall "p"	n.s.	n.s.	n.s.	.001	.001	n.s.	n.s.	n.s.	n.s.

Note. For Confidence and Risk columns the cell entries are the differences in the proportion of responses above the median (higher confidence and higher risk) for individual versus interaction subjects. Positive values indicate that a higher proportion of interaction subjects are above the overall median. Entries in the Change Answers column are differences in the proportion of people choosing the "risky" alternative.

[a] These dilemmas have traditionally shifted conservative.

* Individually significant at $p < .05$.

condition (Wilcoxon $z = 5.2, p < .001$) and the Directed Writing condition (Wilcoxon $z = 4.4, p < .001$), but not the Pro/Con Writing condition.

Treating each prediction as independent, Table 8 reveals that confidence is more likely to go up than down in each condition (all p's $< .01$ under the assumption of independence).

Reactions of deviants. Now we examine the reactions of deviants who were confronted with unanimous opposition of the other subjects with whom they interacted. In interactions where there was a deviant (43 out of the 93 discussions in the experiment), the deviant switched answers to the majority answer 29% of the time, and 15% of the time decreased his or her expressed confidence without changing answers. Both actions might be expected from an information collection perspective since during interaction deviants receive arguments against their initial position. However, the majority of deviants (56%) did not change their answer, and showed stable or increasing confidence in the face of unanimous disagreement—44% kept the same answer with the same confidence, and 12% kept the same answer and actually increased in confidence. As in Study 2, we again report a confidence interval for this proportion because it is not clear how to state the null hypothesis for the test of significance. The 95% confidence interval (41%, 70%) leaves little doubt that deviants exhibit a substantial amount of inertia—almost one of two deviants refused to change their response even in the face of unanimous disagreement.[8]

TABLE 8

Study 3: Within Subjects Analysis—Change in Accuracy by Change in Confidence for Subjects in Three Experimental Conditions

Choice	Confidence: Down	Same	Up	Total
	Interaction condition			
Same	56	150	151	357
Changed	9	7	10	26
Total	65	157	161	
	Directed writing			
Same	24	82	62	168
Changed	1	0	0	1
Total	25	82	62	
	Pro/con writing			
Same	38	100	60	198
Changed	0	3	5	8
Total	38	103	65	

Discussion

Study 3 as a test of information collection. This study provides strong converging evidence for the pattern of results demonstrated in Studies 1 and 2 despite the use of very different experimental materials. As in the Baseline Shift study, the between-subjects test fails to detect any choice shifts after interaction, and the within-subjects tests show that the number of

[8] The 39/43 observations are independent. Two subjects contributed two observations. Omitting data from these two subjects strengthens the current result slightly.

shifts is very small—overall only 6.8% of participants changed their answers in response to the information collected during interaction. If we had merely run the within-subjects condition, the lack of shifts could have been explained by a bolstering or commitment process (Deutsch & Gerard, 1955; Janis & Mann, 1977), but the between-subjects participants never make an initial commitment to any particular choice.

The most surprising result for information collection theories is the stability of deviant choices in the face of unanimous disagreement. In confronting a majority who disagree, subjects should certainly have collected arguments in favor of the other option and against the one they initially favored. After interaction, deviants should change their decisions, or at the very least hold their initial position with less certainty. In fact, the majority of individuals show no change and 12% actually increase in confidence.

There are two alternative ways of looking at the traditional group polarization and risky shift results, and the current results strongly favor one of them:

The labeling of the literature suggests that interaction causes shifts in the *content or extremity* of beliefs—for example, the labels claim that interaction causes choices to "shift" and decisions to "polarize." Some researchers have been explicit about this. At one point, Pruitt (1971) proposed the term "*choice* shift" as a label for the empirical phenomenon. This term has been used frequently ever since (e.g., Burnstein & Vinokur, 1975; Laughlin & Earley, 1982; McGuire, Kiesler, & Siegel, 1987). Based on such descriptions of the phenomenon, we adopted the choice measure as a clean way to examine shifts in "choices" and were surprised not to find much evidence of such shifts.

Although the label suggests some "shift" in the *content* of beliefs, the definition of the risky shift suggests that interaction increases the *strength of beliefs;* that average post-interaction responses on some attitude scale would be more extreme than average individual responses before discussion. This definition is consistent with the current results: Given the stability of choices, the confidence increases in the current study are better interpreted as increases in the *strength* of belief rather than shifts in the *content or extremity* of beliefs. We may become more willing to act on our decisions after interaction (strength of belief), but we are not necessarily willing to take a more radical action (extremity of belief).

Once we clarify that the effects of interaction seem to be more about strength of belief than extremity, we might want to adopt a different conceptual label for studies in these literatures. It may be more reasonable to relabel past demonstrations of "polarization" or "risky shift" as *increases in confidence*. Although it is possible to interpret confidence increases as a kind of polarization, confidence is a more natural psychological construct than "polarization," and lends itself readily to examination in other domains—like the football predictions examined in Studies 1 and 2. In studies that demonstrated "shift" or "polarization," it is reasonable to assume that people became more confident after interaction and used the available response scale (whether a risk or an attitude scale) to signal their increased strength of belief.

Some tentative evidence for the confidence interpretation is given by the failure of the risky shift measure to show any effects in the current studies. After allowing people to signal their confidence using a confidence scale, there was no tendency to respond differently on the risk scale.[9]

Study 3 as a test of rationale construction. Labeling the effects of interaction as "increases in confidence" is consistent with the argument of the rationale construction hypothesis. Although further work needs to be done, the Directed Writing task provides some preliminary evidence that interaction may be better described as rationale construction rather than as information collection. The Directed Writing task succeeded in producing similar confidence increases to the interaction condition by asking people to engage in individual rationale construction without social interaction.

Because the Directed Writing procedure successfully increased confidence while the Pro/Con procedure failed, we can eliminate some other internal processes that could be responsible for confidence increases, for example, familiarity or "mere thought." This is the first time in a long literature that an experiment has successfully mimicked the effects of interaction using a writing task. Previous attempts to model interaction have used writing tasks, which like the Pro/Con task, force subjects to attend to arguments against a hypothesis as well as for it (Bell & Jamieson, 1970; Teger *et al.*, 1970; Burnstein & Vinokur, 1975). If interaction produces its effects by forcing people to create a compelling rationale for their decision, it is not surprising that Pro/Con-style procedures do not produce similar effects to interaction. The Pro/Con procedure prevents rationale construction by encouraging people to attend

[9] Although it might be sufficient in the current study to attribute the failure to find a risky shift to the changed dependent measure, other experience leads us to doubt this interpretation. In five years of experiments within our research group using the traditional risky shift scale we have found that risky shift problems rarely work the way that could be anticipated from reading the literature. In our studies, most problems fail to show the individual effects, and many show effects opposite their traditional shifts (for example, risky problems such as the Company Location problem sometimes shift cautiously). Overall, our research has led us to suspect that the effect is much less robust than the literature would lead us to believe.

to arguments against their choice as well as for it. Forcing people to attend to arguments against their preferred alternative breaks down the rationale for choice instead of building it, and the disjointed nature of this process prevents the formation of a coherent story. Indeed, in the risky shift literature, similar writing conditions sometimes produce "moderation" shifts—risky items show cautious shifts and cautious items show risky shifts (e.g., Teger *et al.*, 1970).

Taken together, these results suggest that interaction may be better modeled as the construction of a coherent case for action than as unbiased consideration of available evidence. By using a plausible model of how information is processed internally, it is possible to create confidence increases similar to those produced by interaction without allowing people to collect any information from the external environment.

GENERAL DISCUSSION

The "information collection" view expressed in Persuasive Arguments Theory and in the interaction survey cannot explain the effects in our studies. Information collection should predict that interaction would improve decisions, however, Studies 1 and 2 indicate that decision accuracy does not improve. Information collection should predict that interaction would increase meta-knowledge, however calibration and resolution do not improve. The simplest test of information collection is that of responsiveness; however, subjects frequently showed stable or increasing confidence even when they interact with others who unanimously disagree with them (Studies 2 and 3).

In contrast to the equivocal effects of interaction on accuracy, meta-knowledge, or responsiveness, Studies 1–3 show that interaction does produce robust increases in confidence. We find increased confidence after interaction using the probability judgments of Studies 1 and 2 and the confidence scale of the Baseline Shift study and Study 3. In addition, subjects in the football studies are more willing to bet on their predictions after interaction. The pattern of unimproved or unresponsive decisions and increased confidence is found for very different experimental materials, providing strong convergent evidence for the pattern's stability.

The phenomenon of confidence increases despite unimproved decisions is not anticipated by existing literature, or by people who are asked to describe their interactions. At the beginning of the article we reported partial results of a survey that asked people to describe their reasons for interaction. We ran the survey after completing the studies in this paper, and knowing our results, we took the opportunity to ask people to directly contrast the information collection view of interaction with the confidence effects demonstrated in these studies. People ranked the confidence-based reasons for interaction (e.g., "helps me organize my thoughts" and "builds confidence in my choice") below the information gathering reasons (e.g., "gives me information to make a better decision" and "causes me to think about things I have not considered") (See Table 9 for ranks). Similar results are obtained if you ask people to describe what happens after interaction instead of asking them why they interact—people believe that interaction is successful at doing what they intend for it to do.

When Will People Respond to Information?

The current results do not suggest the extreme position that people never collect information from their interactions. It is unlikely, for example, that someone with a broken transmission but little knowledge of cars would ignore the advice of an honest and impartial car repair expert (should one be found).

The current studies contain some hints about when information collection may play a relatively more important role in interaction. Information collection probably plays a bigger role when people lack knowledge of a domain and when they believe that the available information is important for making good decisions. Thus, it is unsurprising that more deviants changed their choices on the football predictions in Study 2 than in the choice dilemmas of Study 3. People may also be more likely to respond to information when people have an ongoing relationship with their interaction partners (as much to preserve social harmony as because they have greater faith in the validity of information). The football partners of Study 1 were more likely to modify their answers when they disagreed than the deviants of Study 2 and Study 3 who did not have an ongoing relationship with their interaction partners.

However, questioning whether information collection ever has *some* effect is really beside the point. The

TABLE 9

Survey of (n = 118) Stanford Students Assessing Their Reasons for Interacting with Others on Important Life Decisions

Avg. rank	
2.42	Gives me information to make a better decision.
2.47	Causes me to think about things I have not considered.
3.03	Helps me organize my thoughts.
3.56	Builds my confidence in my decision.
4.49	Helps me avoid making mistakes.
4.81	Provides social support.

Note. Subjects ranked the six reasons listed from 1 = 'most important' to 6 = 'least important.'

question addressed by our studies is more a question of effect magnitude than a question of possibility. It is much harder to find evidence of information collection in the current studies than we would have expected before we began this program of research, and these results have surprised us.

Rationale Construction as an Explanation for Interaction

Studies 1–3 and the Baseline Shift study show that subjects show consistent and robust confidence increases after interaction. These confidence increases occur despite a lack of corresponding evidence that the information collected during interaction successfully improves or modifies decisions. The Directed Writing task of Study 3 produced similar confidence increases without giving subjects any opportunity to collect information.

To explain these confidence increases, we rejected explanations that assumed that confidence increases happened because people perceived themselves as having collected information from the interaction (perceived information gain) or because they perceived themselves as having their opinions validated by others (inference certification). Instead, we proposed that confidence increases result from an internal process that is prompted by the social environment but that does not require information from the social environment. The *rationale construction* view of interaction says that when people interact, they are forced to better organize and elaborate their case in favor of their choice, and this process leads them to become more confident in their decisions. A large number of studies in social psychology have found that individuals show increased confidence after explaining their predictions and judgments (cf. Anderson & Wright, 1988; Koehler, 1991). The literature on explanation generation extends the literature on interaction because it describes a psychological process that can explain confidence increases during interaction. In turn, the literature on interaction extends the literature on explanation generation because it shows that there is an ecologically important context which forces people to engage in explanation generation.

Implications of the Current Research

The current results have methodological and theoretical implications. On the methodological side, these results emphasize the advantage of using interactive decision procedures to study decisions in a social environment. The most interesting effects of the current research—deviants' stability in choices and confidence—could not have been detected using consensual group responses.

On a theoretical level, distinguishing between interactive and group procedures may also expand our understanding of social interaction. Rationale construction may provide an example of an interesting psychological difference between interactive and group procedures. When we are in groups that must come to a consensual decision, we may spend more time looking for areas of agreement than trying to explain our own position, and in turn we may be more open to the information of others. This implies that rationale construction may be more prevalant in interactive decisions than in group decisions—when individuals think they must work towards a consensual decision, they may be more willing to modify their decisions on the basis of the information they collect from others.

Also on the theoretical side, these results might question when we should argue that a phenomenon is based on group dynamics as opposed to social dynamics. In his classic writings on "groupthink" (Janis, 1971), Janis argues that instead of thinking critically, members of cohesive, long-term groups often internalize the values of the group to such an extent that they avoid considering information in a balanced way. "In a cohesive group, the danger is that each individual . . . will think the proposal is a good one without attempting to carry out a careful, critical scrutiny of the pros and cons of the alternatives." Because of their cohesion, groups often "discount warnings and other forms of negative feedback that, taken seriously, might lead the group members to reconsider their assumptions." Stripped of their emphasis on "groups," these quotes provide a remarkably satisfying description of the phenomenon we describe in this paper, which is a product not of cohesive, long-term groups, but of individual decision makers interacting in a social environment. The current results indicate that "groups" are not necessary to produce groupthink-style outcomes.

Confidence in a Social Environment

Even if interaction does not increase the quality of our decisions, it may play an important social role by increasing decision confidence. Psychologists have spent the last few years exploring the adaptiveness of "positive illusions" (e.g., Taylor & Brown, 1988). In a world that more often rewards action than inaction, people who interact to increase their confidence and then take some (any?) action may succeed more often than those who remain wrapped in doubt and refuse to act. Eisenhardt (1989) argues that interaction serves as a confidence-builder for top Silicon Valley executives who are regularly forced to act quickly in a volatile, high-stakes environment.

We can view information collection and rationale construction as two different methods of reducing the uncertainty associated with difficult decisions. Information collection theories assume that decision uncertainty is located in the *external environment*—people interact because they need to fill in gaps in their knowledge of the world, clarify ambiguous signals, or improve their predictions about the world. Rationale construction assumes that decision uncertainty is located *internally*—people interact because they need to fill in gaps in their justification for their decision, clarify contradictory impulses, or integrate conflicting information. In the process of explaining their thoughts and confusions to others, people develop a clearer, more coherent, and more compelling rationale for their actions, and this in turn may make them more willing to take action on their beliefs. Thus this line of research has the potential to address some interesting questions beyond those about social dynamics. By exploring why interaction helps us make difficult decisions, we may better understand why decisions are difficult.

Instead of Ambrose Bierce's cynical definition of conversation that we reported in the introduction, we close with an observation by Benjamin Franklin. Like Bierce, Franklin proposed that interaction helps us "arrange our wares," but he describes interaction more positively: "Reading makes a full man, meditation a profound man, discourse a clear man." Even when interaction does not change our minds or produce more accurate choices, it may allow us to clarify our thoughts about a decision and move forward with confidence. In a world in which we are often uncertain about our ability to choose wisely, that increased confidence may allow us to continue to act.

APPENDIX

Choice Dilemmas for Studies 3 and 4

Below are the problems used in Studies 3 and 4. The Med School and Student problems are new. The other six problems are only slightly modified from the Kogan and Wallach (1967) Choice Dilemmas questionnaire (e.g., names have been added rather than designating the characters as Mr. A, Mr. C, etc.).

(**Surgeon**) Dr. Hughes is a thoracic surgeon with a well-established surgical practice. He is married and has three children, one of which is just starting college. During a backyard session of football, Dr. Hughes seriously dislocated his shoulder. Although the shoulder was properly reset at the time, the dislocation produced some nerve damage, and he has been experiencing a great deal of pain ever since. An operation is available that will relieve the pain if completely successful, but the operation also poses a risk of producing a permanent decrement in manual dexterity. The decrement in dexterity is normally inconsequential, but in Dr. Hughes' case, it could prevent him from continuing his surgical practice.

(**Vacationer**) Mr. Baker is about to board a plane at the airport to begin an overseas vacation. Although he has been looking forward to this trip for some time, he is troubled because he awoke this morning with a severe abdominal pain. Because he has never flown before, he thinks that the pain may simply be an upset stomach brought on by anticipation of his flight. Although he is not far from a hospital where he knows he will obtain quick attention, he realizes that a visit to the hospital will cause him to miss his flight. This in turn will seriously disrupt his vacation plans. The pain has gotten more severe in the last few minutes.

(**Football**) Dave is the captain of College X's football team. College X is playing an important conference opponent in the final game of the season. The game is in its final seconds, and Dave's team is behind in the score. College X has time to run one more play. Dave is considering two plays. One play, very routine and cautious, would be certain to end the game in a tie. On the other hand, he could try a more complicated and risky play which could bring victory if it was successful, but defeat if not.

(**Company Location**) Mr. E is president of a light metals corporation in the United States. The corporation is quite prosperous, and has strongly considered the possibilities of business expansion by building an additional plant in a new location. The choice is between building another plant in the U.S. where there would be a moderate return on the initial investment, or building a plant in a foreign country. Lower labor costs and easy access to raw materials in that country would mean a much higher return on the initial investment. On the other hand, there is a history of political instability and revolution in the country under consideration. In fact, the leader of a small minority party is committed to nationalizing (taking over) all foreign investments.

(**Med school**) Elizabeth is a senior majoring in biology. All her life she has wanted to practice medicine. In college she has maintained a good grade point average. She works part-time to supplement her scholarship because her parents are unable to provide financial assistance.

Two medical schools have offered Elizabeth admission. Both schools are private schools, so the tuition and other costs are relatively comparable.

- School A is one of the best medical schools in the country. Both the faculty and the technical resources of the school rank it among the top 10 in the nation. Unfortunately, school A will not offer more than 20% of tuition in financial aid. Therefore, Elizabeth would have to take out large loans to pay for her education. (20% aid)

- School B does not have as prestigious a faculty, but is respectable and provides a good medical education. It has offered Elizabeth a comprehensive financial aid package which would pay for 75% of her tuition. (75% aid)

Imagine that you are advising Elizabeth on her medical school decision. The decision of which medical school to attend is very complex and depends on many factors. Take a few moments to consider which of the two alternatives you would recommend.

(**Engineer**) Andrew Whittaker, an electrical engineer is married and has one child, has been working for a large electronics corporation since graduating from college. He is assured of a lifetime job with a modest though adequate salary, and liberal pension benefits upon retirement. On the other hand, it is very unlikely that his salary will increase tremendously before he retires. While attending a convention, Andrew is offered a job with a small, newly founded company which has a highly uncertain future. The new job would pay more to start and would offer the possibility of a share in the ownership of the company if it survived the competition of the market.

(**Student**) Sam G. has had a difficult fall semester carrying 18 upperdivision units and working part time. Through the semester, his "light at the end of the tunnel" has been a ski trip in Aspen. He has been planning the trip since the first week of the trimester.

Sam has finished two of his 5 finals—he is quite exhausted and is looking forward to enjoying the upcoming ski trip with the college group. The night before his third final he gets a phone call from his best friend from home, whose parents have suddenly divorced. His friend is really hurting and quite confused. After the phone call Sam wonders whether he should skip the ski trip and visit his friend instead. Whatever the case, he would see his friend in three weeks on his next visit home.

(**Career**) Roger Wilkins, a college senior, has studied the piano since childhood. He has won amateur prizes and given small recitals, suggesting that Roger has considerable musical talent. As graduation approaches, Roger has the choice of taking a medical school scholarship to become a physician, a profession which would bring certain financial rewards, or entering a conservatory of music for advanced training with a well-known pianist. Roger realizes that even upon completion of his piano studies success as a concert pianist would not be assured.

REFERENCES

Abelson, R. (1973). Comment on 'Group shift to caution at the race track:. *American Psychologist*, 9, 517–521.

Anderson, C. A. (1983). Imagination and expectation: The effect of imagining behavioral scripts on personal intentions. *Journal of Personality and Social Psychology*, 45, 293–305.

Anderson, U., & Wright, W. F. (1988). Expertise and the explanaton effect. *Organizational Behavior and Human Decision Processes*, 42, 250–269.

Baron, R. S., & Roper, G. (1976). Reaffirmation of social comparison views of choice shifts: Averaging and extremity effects in an autokinetic situation. *Journal of Personality and Social Psychology*, 33, 521–530.

Bell, P. R., & Jamieson, B. D. (1970). Publicity of initial decisions and the risky shift phenomenon. *Journal of Experimental Social Psychology*, 6, 329–345.

Bishop, G. D., & Myers, D. G. (1974). Informational influence in group discussion. *Organizational Behavior and Human Performance*. 12, 92–104.

Brown, R. (1965). Further comment on the risky-shift. *American Psychologist*, 29, 468–470.

Burnstein, E. (1982). Persuasion as argument processing. In H. Brandstatter, J. H. Davis, & G. Stocher-Kreichgauer (Eds.), *Contemporary problems in group decision-making* (pp. 103 –124). New York: Academic Press.

Burnstein, E., & Vinokur, A. (1973). Testing two classes of theories about group-induced shifts in individual choice. *Journal of Experimental Social Psychology*, 9, 123–137.

Burnstein, E., & Vinokur, A. (1975). What a person thinks upon learning he has chosen differently from others: Nice evidence for the persuasive arguments explanation of choice shifts. *Journal of Experimental Social Psychology*, 11, 412–426.

Burnstein, E., & Vinokur, A. (1978). Depolarization of attitudes in groups. *Journal of Personality and Social Psychology*, 36, 872–885.

Campbell, J. D., & Fairey, P. J. (1985). Effects of self-esteem, hypothetical explanations, and verbalization of expectancies on future performance. *Journal of Personality and Social Psychology*, 48, 1097–1111.

Cartwright, D. (1971). Risk taking by individuals and groups: An assessment of research employing choice dilemmas. *Journal of Personality and Social Psychology*, 20, 361–378.

Cochran, W. G. (1954). Some methods for strengthening the common χ^2 test. *Biometrica*, 10, 417–451.

Davis, J. H. (1973). Group decision and social interaction: A theory of social decision schemes. *Psychological Review*, 80, 97–125.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgement. *Journal of Abnormal and Social Psychology*, 51, 629–636.

Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology*, 58(4), 568–581.

Dunning, D., & Ross, L. (1985). *Overconfidence in individual and group prediction: Is the collective any wiser?* Unpublished manuscript, Stanford University.

Ebbesen, E. B., & Bowers, R. J. (1974). Proportion of risky and conservative arguments in a group discussion. *Journal of Personality and Social Psychology*, 29, 316–327.

Eisenhardt, K. M. (1989). Making fast strategic decisions in high-velocity environments. *Academy of Management Journal*, 32, 543–576.

Fazio, R. H. (1979). Motives for social comparison: The construction-validation distinction. *Journal of Personality and Social Psychology*, 37, 1683–1698.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of*

Experimental Psychology: Human Perception and Performance, 3, 552–564.

Hastie, R. (1986). Experimental evidence on group accuracy. In B. Grofman & G. Owen (Eds.), *Decision research* (Vol 2, pp. 129–157). Greenwich, CT: JAI Press.

Hill, G. W. (1982). Group versus individual performance: Are $N + 1$ heads better than one? *Psychological Bulletin,* **91**, 517–539.

Heath, C., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty,* **4**, 5–28.

Hoch, S. J. (1985). Counterfactual reasoning and accuracy in predicting personal events. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **11**, 719–731.

Isenberg, D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology,* **50**, 1141–1151.

Janis, I. L. (1971, November). Groupthink. *Psychology Today Magazine.*

Janis, I. L. (1972). *Victims of groupthink.* Boston: Houghton Mifflin.

Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment.* New York: The Free Press.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases.* Cambridge, England: Cambridge Univ. Press.

Kaplan, M. F. (1977). Discussion polarization effects in a modified jury decision paradigm: Informational influences. *Sociometry,* **40**, 262–271.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgement. *Psychological Bulletin,* **110**, 499–519.

Kogan, N., & Wallach, M. A. (1967). The risky-shift phenomenon in small decision making groups: A test of the information-exchange hypothesis. *Journal of Experimental Social Psychology,* 3, 75–85.

Kogan, N., & Wallach, M. A. (1967b). Effects of physical separation of group members on group risk taking. *Human Relations,* **20**, 41–48.

Lamm, H. (1967). Will an observer advise higher risk taking after hearing a discussion of the decision problem? *Journal of Personality and Social Psychology,* **6**, 467–471.

Laughlin, P. R., & Earley, P. C. (1982). Social combination models, persuasive arguments theory, social comparison theory, and choice shifts. *Journal of Personality and Social Psychology,* 42, 273–280.

Liberman, V., & Tversky, A. (1993). On the evaluation of probability judgements: Calibration, resolution, and monotonicity. *Psychological Bulletin,* **114**.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1980). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds). *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). Cambridge: Cambridge Univ. Press.

McGuire, T. W., Kiesler, S., & Siegel, J. (1987). Group and computer-mediated discussion effects in risk decision making. *Journal of Personality and Social Psychology,* 52, 917–930.

Received: October 20, 1993

Myers, D. G., & Bishop, G. D. (1971). The enhancement of dominant attitudes in group discussion. *Journal of Personality and Social Psychology,* **20**, 386–391.

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin,* **83**, 602–627.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgement.* Englewood Cliffs: Prentice–Hall.

Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the story model for juror decision making. *Journal of Personality and Social Psychology,* **62**, 189–206.

Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgement. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **14**, 521–533.

Pruitt, D. G. (1971). Choice shifts in group discussion: An introductory review. *Journal of Personality and Social Psychology,* **20**, 339–360.

Ross, L., & Anderson, C. A. (1982). Shortcomings in the attribution process: On the origins and maintenance of erroneous social assessments. In D. Kahneman, P. Slovic, & A. Tversky (Eds). *Judgment under uncertainty: Heuristics and biases* (pp. 127–152). Cambridge: Cambridge Univ. Press.

Sherman, S. J., Skov, R. B., Hervitz, E. F., & Stock, C. B. (1981). The effects of explaining hypothetical future events: From possibility to probability to actuality and beyond. *Journal of Experimental Social Psychology,* 17, 142–158.

Sherman, S. J., Zehner, K. S., Johnson, J., & Hirt, E. R. (1983). Social explanation: The role of timing, set and recall on subjective likelihood estimates. *Journal of Personality and Social Psychology,* **44**, 1127–1143.

Sniezek, J. A., & Buckley, T. Cueing and cognitive conflict in Judge–Advisor decision-making. *Organizational Behavior and Human Decision Processes,* in press.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin,* **103**, 193–210.

Teger, A. I., Pruitt, D. G., St. Jean, R., & Haaland, G. A. (1970). A reexamination of the familiarization hypothesis in group risk taking. *Journal of Experimental Social Psychology,* **6**, 346–350.

Tesser, A. (1978). Self-generated attitude change. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 11). New York: Academic Press.

Vinokur, A., & Burnstein, E. (1978). Depolarization of attitudes in groups. *Journal of Personality and Social Psychology,* **36**, 872–885.

von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research.* Cambridge, England: Cambridge Univ. Press.

Yaniv, I., Yates, J. F., & Smith, J. E. K. (1991). Measures of discrimination skill and probability judgments. *Psychological Bulletin,* **110**, 611–617.

Yates, J. F. (1991). JDM conference presentation, November, San Francisco.