Social Hypothesis Testing
and the Role of Expertise

Eugene Borgida
University of Minnesota

Kenneth G. DeBono
Union College

Research on social hypothesis-testing processes has demonstrated that there is a strong preference for seeking and using hypothesis-confirming evidence when testing hypotheses about other people. An experiment was conducted to examine the extent to which expertise in a particular knowledge domain represents a delimiting condition of this inferential tendency. Expert and novice subjects completed a hypothesis-testing recall task. Recall differences in line with the different strategies commonly used by experts and novices on problem-solving and reasoning tasks were predicted. As expected, novices recalled more hypothesis-matching evidence after a delay interval, regardless of the nature of the hypothesis to be tested. Expert recall, by contrast, was not characterized by this confirmatory preference when experts tested the hypothesis inside their domain of expertise. The extent to which these expert/novice differences provide support for motivated social cognition is discussed.

Past research in social cognition has demonstrated that, under certain conditions, there is a preference for seeking and using hypothesis-confirming information when testing hypotheses about other people (Fiske & Taylor, 1984; Sherman & Corty, 1984; Snyder & Gangestad, 1981). In the first demonstration of this phenomenon, Snyder and Swann (1978) asked subjects to test one of two hypotheses about a target person: either that the target was an extraverted individual or that the target was an introverted individual. Subjects were allowed to select a series of questions they believed they would be asking the target person in an upcoming getting-acquainted session. Snyder and Swann found that subjects preferentially selected those questions that would elicit hypothesis-confirming responses. That is, if the hypothesis was that the target was an extraverted individual, subjects chose questions such as "What would

you do if you wanted to liven things up at a party?" If the hypothesis was that the target was an introverted individual, subjects chose questions such as "What factors make it hard for you to open up to other people?"

Additional support for this confirmatory preference comes from a study by Snyder and Cantor (1979), who had subjects read about 1 week in the life of a woman named Jane. The story about Jane was designed so that Jane, in the course of the week, acted in an approximately equal number of introverted and extraverted ways. Two days after reading the story, subjects returned and were told either that Jane was applying for the job of archivist and research librarian and that they were to determine her suitability for this rather introverted position or that Jane was applying for the job of real estate salesperson and that they were to determine her suitability for this rather extraverted position. Subjects were then asked to recall all the facts from the story that they felt were relevant to their decisions regarding Jane's suitability for the job in question.

Snyder and Cantor (1979) found that subjects in both conditions tended to recall more hypothesis-confirming than hypothesis-disconfirming information. That is, subjects who were told that Jane was applying for the job of archivist and research librarian tended to recall more instances of Jane's acting in an introverted manner than in an extroverted manner. By contrast, subjects who were told that Jane was applying for the job of real estate salesperson tended to recall more instances of Jane's acting in an extraverted manner than in an introverted manner.

Trope and colleagues (Bassok & Trope, 1984; Trope & Bassok, 1982, 1983; Trope, Bassok, & Alon, 1984; Trope & Mackie, 1987; see also Skov & Sherman, 1986), however, have recently challenged the contention that individuals employ a hypothesis-confirming strategy when testing hypotheses about others or remembering information about others. In a series of investigations, they have convincingly demonstrated that, when testing a hypothesis about another's personality and given the choice of information to use, individuals do not selectively choose evidence that will confirm the hypothesis but instead select evidence that is highly diagnostic of the personality trait under scrutiny. Thus they contend that the preference for confirmatory evidence found in the hypothesis-testing task used by Snyder and his colleagues (e.g., Snyder, 1981) may be real, but clearly limited to certain kinds of question-asking tasks. As Higgins and Bargh (1987) have succinctly observed, "When given the opportunity, people select diagnostic questions over nondiagnostic questions (i.e., they 'seek the truth') but will display a preference for hypothesis-matching questions when the hypothesis under consideration is the only accessible alternative and is believed to be true" (pp. 401-402).

The preference for confirmatory evidence may also be limited by individual difference considerations. In this case, expert knowledge in the domain of the hypothesis to be tested may determine when a preference for hypothesis-confirming information will be found. Expertise, by definition, represents knowledge and firsthand experience in a particular domain (e.g., Chase &

Simon, 1973; Larkin, McDermott, Simon, & Simon, 1980). Not only do experts have more knowledge and experience in a given domain, but their knowledge may be structured differently than that of novices. In contrast to the knowledge structure of novices, the knowledge structure of experts tends to be more differentiated (Johnson, Hassebrock, Duran, & Moller, 1982), more tightly organized (McKeithen, Reitman, Rueter, & Hirtle, 1981; Chi & Koeske, 1983), and more hierarchically structured (Johnson, Hassebrock, Duran, & Moller, 1982).

In addition to differences between experts' and novices' knowledge structures, experts and novices differ in how their knowledge is used to solve problems in a given area of expertise. Experts in many different fields tend to use more efficient, sophisticated, and efficacious strategies to solve problems and make inferences in their domain of expertise (Fiske & Kinder, 1981). Of particular interest are studies demonstrating that experts tend to be more sensitive to information that is inconsistent with their prior knowledge of the domain (Fiske, Kinder, & Larter, 1983; Johnson, Hassebrock, Duran, & Moller, 1982). This research demonstrates not only that experts differ from novices in the contents of their knowledge but also that they tend to use their knowledge in a different fashion.

To what extent, therefore, does expertise affect the preference for and recall of hypothesis-confirming evidence? More specifically, when experts test a hypothesis about another person in their domain of expertise, is it the case that they will be *less* likely to seek out hypothesis-confirming information than novices? Experts, we would expect, should be more sensitive to and more likely to use hypothesis-disconfirming information. When testing hypotheses *outside* their domain of expertise, however, where experts no longer are experts, we would expect experts to behave like novices and to better remember hypothesis-confirming information.

METHOD

Subjects

Novice sample. Twenty male and female undergraduates enrolled in psychology classes at the University of Minnesota volunteered to participate in the study.

Expert sample. Twenty members of the library staff at the University of Minnesota, all of whom were full-time professional librarians, volunteered to participate in the study.

Procedure

Session 1. The procedure in this session (as well as in the second session) was based on the procedure used by Snyder and Cantor (1979, Study 2). Subjects in both samples first received a five-page booklet describing 1 week in the life of an ostensibly real person named Jane. They were instructed that, while reading the

story, they should concentrate on the factual details contained in the story because, they were told, this research was investigating memory for detailed information. The story was identical to the one used in the Snyder and Cantor study. The protocol had been pretested by those investigators such that (a) Jane acted in an approximately equal number of introverted and extraverted ways and (b) Jane was viewed by judges as being both moderately introverted and moderately extraverted. All subjects were given 7 minutes to complete reading the story.

Session 2, novice sample. All subjects in the novice sample returned to the lab 2 days later. Upon returning, they were randomly assigned to one of two conditions: either to the real estate salesperson condition or to the research librarian condition. Those in the *real estate salesperson condition* received a booklet saying that Jane was applying for the job of real estate salesperson in her neighborhood and that they were going to judge her suitability for this position. Subjects assigned to the *research librarian condition* received a similar booklet, but this version informed them that Jane was applying for the job of research librarian at the local college library and that they were going to judge her suitability for this position.

After reading these instructions, subjects were given 7 minutes to write down all those facts they could recall about Jane that they considered relevant to their decision concerning Jane's suitability for the job in question. After completing the recall measure, subjects first were instructed to indicate which of the facts they recalled would make Jane suitable for the job (thus confirming the hypothesis), which facts would make her not suitable (thus disconfirming the hypothesis), and which facts would have no impact.[1] When they finished this task, subjects were then asked to indicate Jane's suitability for the job considered on a 6-point scale (1 = *extremely unsuited*, 6 = *extremely suited*).

Session 2, expert sample. After a similar delay, subjects in the expert sample returned and, like subjects in the novice sample, were randomly assigned to either the real estate salesperson condition or the research librarian condition. The procedure and the stimulus materials used for the expert sample were exactly the same as for the novice sample.

RESULTS

Recall Analysis

To derive an index for assessing the type of hypothesis-testing strategy used, for each subject we subtracted the number of self-scored disconfirming items recalled from the number of self-scored confirming items. Positive scores on this index indicate that a preponderance of hypothesis-confirming items were recalled, and negative scores on this index indicate that relatively more hypothesis-disconfirming items were recalled. Table 1 presents the mean difference scores.

To test our prediction, a planned comparison was performed (Hays, 1973). Only experts testing a hypothesis in their domain of expertise were expected to

TABLE 1   Mean Confirming–Disconfirming Facts
          Recalled by Experts and Novices

| | Hypothesis Tested | |
Subjects	Real Estate Salesperson	Research Librarian
Experts		
M	2.9	−1.2
SD	4.04	2.93
n	10	10
Novices		
M	2.5	3.0
SD	2.27	3.02
n	10	10

Note: Positive scores indicate that relatively more hypothesis-matching items were recalled. Negative scores indicate that relatively more hypothesis-disconfirming items were recalled.

recall comparable amounts of confirming and disconfirming evidence. A weight of -3, therefore, was assigned to the difference score (confirmatory–disconfirmatory) of the expert sample / research librarian condition, and weights of +1 were assigned to the difference scores of the other three conditions. The ANOVA contrast on confirmatory minus disconfirmatory difference scores was highly reliable, $F(1, 36) = 7.02, p < .001$, suggesting that when the hypothesis to be tested is outside their field of expertise (in this case, real estate), experts behave just like novices and recall more hypothesis-confirming evidence from the Jane story. But when the hypothesis is within their field of expertise (library science), this pattern disappears and the experts recall comparable amounts of hypothesis-confirming and hypothesis-disconfirming evidence. Consistent with this pattern, experts testing the real estate hypothesis indeed remembered more hypothesis-confirming facts than experts testing the research librarian hypothesis, $F(1, 36) = 8.34, p < .001$. Novices testing the real estate hypothesis did not differ from novices testing the librarian hypothesis, $F(1, 36) < 1, ns$. Moreover, whereas post hoc analyses using Duncan's multiple range test suggest that experts and novices did not differ in the number of *confirming* items recalled for either real estate salesperson or research librarian, these analyses also show that the difference between the number of *disconfirming* items for experts and novices for research librarian was significant ($p < .05$).

Suitability Ratings

Table 2 presents the mean suitability ratings. Since only experts testing the hypotheses within their domain of expertise were not expected to rate Jane as relatively suitable for the job under consideration, the following planned

TABLE 2   Mean Suitability Ratings by Experts and Novices

| | Hypothesis Tested | |
Subjects	Real Estate Salesperson	Research Librarian
Experts		
M	3.7	2.6
SD	1.06	1.43
n	10	10
Novices		
M	4.4	4.2
SD	1.07	.79
n	10	10

Note: Higher scores indicate greater perceived suitability.

comparison was performed. A weight of –3 was assigned to the mean suitability rating of the expert sample / research librarian condition, and weights of +1 were assigned to the means in the other three conditions. This contrast was highly significant, $F(1, 36) = 7.94, p < .001$, indicating that, as predicted, experts tended not to view Jane as suitable when testing hypotheses within their field of expertise. Consistent with this finding, experts testing the real estate hypothesis viewed Jane as more suitable than experts testing the librarian hypothesis, $F(1, 36) = 5.65, p < .001$. By contrast, novices testing the real estate hypothesis did not view Jane as more suitable than novices testing the librarian hypothesis, $F(1, 36) < 1, ns$.

DISCUSSION

The present research examined the extent to which expert knowledge in a given domain influences the use of a confirmatory strategy in rather constrained hypothesis-testing situations. As predicted, it was demonstrated that experts do *not* preferentially remember hypothesis-confirming information when testing a hypothesis in their domain of expertise. When testing hypotheses in their field of expertise, experts were more sensitive to information that could potentially disconfirm the hypothesis in question. By contrast, when experts tested a hypothesis outside their domain of expertise, they behaved much as the novices did and preferentially remembered hypothesis-confirming information.

The differences uncovered between expert and novice hypothesis-testing strategies appear to be a function of differences in the strategies that experts and novices may have used when processing information. For example, it may be the case that the intuitive strategy employed when testing a hypothesis in a task situation like ours resembles the strategy used by individuals when assessing the degree of covariation between two variables. Research on covariation detection

has suggested that when individuals are presented with a 2 × 2 contingency table, they tend to rely primarily on the content of the present-present cell (i.e., the number of times Event 1 and Event 2 co-occurred) to arrive at their judgments (Arkes & Harkness, 1983; Alloy & Tabachnik, 1984).

This simple confirmatory strategy, or "positive test strategy" (Klayman & Ha, 1987), so prevalent across a wide variety of covariation detection tasks, has also been implicated in hypothesis-testing situations (Snyder, 1981; Sherman & Corty, 1984). Consequently, the information that we have been labeling "hypothesis-confirming" may simply reflect the number of times that two variables co-occur. In our research, subjects were asked whether there was any relationship between Jane (Event 1) and the job of research librarian/real estate salesperson (Event 2). Additionally, subjects were free to use whatever type of information they felt was "relevant to their decision regarding Jane's suitability" for this job. Consistent with the use of a "positive test strategy," subjects tended to recall those instances when Jane's attributes overlapped with or "matched" the attributes of a research librarian; that is, they recalled the times when the event "Jane" and the event "research librarian" co-occurred.

Experts, by contrast, do not appear to use such a simple strategy when making judgments in their domain of expertise. Experts seem able to use their knowledge in a more sophisticated manner. This differential use of knowledge is manifest as an increased sensitivity to instances when two events do *not* co-occur. Experts, therefore, should be able to recognize the importance of the times when Event 1 (Jane) and Event 2 (research librarian) were not both present. In particular, they were sensitive to the instances when Event 1 (Jane) did occur but Event 2 (research librarian) did not occur. Experts, by implication, should be able to generate alternative competing hypotheses more easily than novices, which should also decrease their preference for hypothesis-confirming information (Higgins & Bargh, 1987).

In the present research, perhaps because experts were more aware of the importance of alternative hypotheses, they were more likely to recall disconfirming evidence when asked whether Jane and the job of research librarian were related. It would be of interest to determine whether, when actually judging the degree of covariation between two events in their domain of expertise, experts also rely primarily on the number of entries in the present/present cell or whether they also take into account the information contained in the other three cells of a covariation matrix. Given the results of our research, we would expect that they would indeed be more sensitive to information contained in these other cells and therefore would be more likely to use a complex strategy (see Harkness, DeBono, & Borgida, 1985).

Such an analysis of the role of individual differences like expertise in hypothesis-testing contexts generally supports the view associated with "motivated social cognition" that people are capable of displaying a motivated, strategic flexibility in their social information processing (Cantor & Kihlstrom, 1981; Markus & Zajonc, 1985; Showers & Cantor, 1985). Reasoning and

memory processes as they affect the interpretation of situations and the planning of behavior generally may be quite responsive to a variety of individual differences, goals, and levels of personal involvement and expertise (Srull & Wyer, 1986). For example, it may have been the case in the present research that librarians (experts), who apparently have a rather sophisticated and extraverted self-image (Wilson, 1981), were motivated to use a hypothesis-disconfirming strategy because the task may have activated self-image considerations. To the extent that librarians regard themselves as extraverted, they may have been more likely to attend to those introverted items that were distinctive or inconsistent with their self-image. However, if our librarians were in fact more introverted, then they may have paid more attention to and better remembered extraverted items. In either case, a mixture of confirming and disconfirming items would be recalled as a result.

Conclusions about covariation detection (Harkness, DeBono, & Borgida, 1985) or, for example, people's use of base-rate information (Borgida & Brekke, 1981) or the extent to which people engage in "top of the head" processing (Borgida & Howard-Pitney, 1983), therefore, must be tempered by a consideration of pertinent task variables and individual differences. The domain of hypothesis testing is no exception, as the present research would suggest. The controversy surrounding this task situation, a controversy that fundamentally questions the appropriateness of labeling certain tasks as hypothesis testing, may be more apparent than real if one considers that social cognition may be strategically motivated. As Markus and Zajonc (1985) and others (e.g., Srull & Wyer, 1986) have suggested, which strategy (in this case, diagnostic versus hypothesis-confirming) is more apt may depend, to a great extent, on the goals of the information search under investigation. Thus, different task goals may lead to the use of different cognitive strategies, and there will therefore be times when people use a diagnostic strategy and times when people use a hypothesis-confirming strategy. Strong support in the present research for the greater flexibility of experts than novices in a hypothesis-testing situation is yet another reason that social cognition research should continue to investigate the role of individual differences in complex cognitive situations.

NOTE

[1] This self-coding procedure was preferred to a procedure involving interrater reliability. The strength of the former procedure is that the experts themselves classify recalled items. In this case, having to instruct judges to make suitability ratings in the absence of any empirical knowledge about the nature of expert stereotypes would render the latter approach less feasible.

REFERENCES

Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review, 91*, 112-149.

Arkes, H. R., & Harkness, A. (1983) Estimates of contingency between two dichotomous variables. *Journal of Experimental Psychology: General, 112*, 117-135.

Bassok, M., & Trope, Y. (1984). People's strategies for testing hypotheses about another's personality: Confirmatory or diagnostic? *Social Cognition, 2*, 199-213.

Borgida, E., & Brekke, N. (1981). The base rate fallacy in attribution and prediction. In J. H. Harvey, W. Ickes, & R. F. Kidd (Eds.). *New directions in attribution research* (Vol. 3). Hillsdale, NJ: Erlbaum.

Borgida, E., & Howard-Pitney, B. (1983). Personal involvement and the robustness of perceptual salience effects. *Journal of Personality and Social Psychology, 45*, 560-570.

Cantor, N., & Kihlstrom, J. F. (Eds.). (1981). *Personality, cognition, and social interaction*. Hillsdale, NJ: Erlbaum.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.

Chi, M.T.H., & Koeske, R. (1983). Network representation of a child's dinosaur knowledge. *Developmental Psychology, 19*, 29-39.

Fiske, S. T., & Kinder, D. T. (1981). Involvement, expertise, and schema use: Evidence from political cognition. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality, cognition, and social interaction*. Hillsdale, NJ: Erlbaum.

Fiske, S. T., Kinder, D. T., & Larter, W. M. (1983). The novice and the expert: Knowledge-based strategies in political cognition. *Journal of Experimental Social Psychology, 4*, 381-400.

Fiske, S. T., & Taylor, S. E. (1984). *Social cognition*. Reading, MA: Addison-Wesley.

Harkness, A. R., DeBono, K. G., & Borgida, E. (1985). Personal involvement and strategies for making contingency judgments: A stake in the dating game makes a difference. *Journal of Personality and Social Psychology, 49*, 22-32.

Hays, W. L. (1973). *Statistics for the social sciences*. New York: Holt, Rinehart and Winston.

Higgins, E. T., & Bargh, J. A. (1987). Social cognition and social perception. *Annual Review of Psychology, 38*, 369-426.

Johnson, P. E., Hassebrock, F., Duran, A. S., & Moller, J. H. (1982). Multimethod study of clinical judgment. *Organizational Behavior and Human Performance, 30*, 201-230.

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

Larkin, J., McDermott, J., Simon, D., & Simon, H. (1980). Expert and novice performance in solving physics problems. *Science, 20*, 1335-1342.

Markus, H., & Zajonc, R. B. (1985). The cognitive perspective in social psychology. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed.). Reading, MA: Addison-Wesley.

McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology, 13*, 307-325.

Sherman, S. J., & Corty, E. (1984). Cognitive heuristics. In R. S. Wyer & T. Srull (Eds.), *Handbook of social cognition* (Vol. 1. pp. 189-286). Hillsdale, NJ: Erlbaum.

Showers, C., & Cantor, N. (1985). Social cognition: A look at motivated strategies. *Annual Review of Psychology, 36*, 275-306.

Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology, 22*, 93-121.

Snyder, M. (1981). Seek, and ye shall find: Testing hypotheses about other people. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario Symposium* (Vol. 1). Hillsdale, NJ: Erlbaum.

Snyder, M., & Cantor, N. (1979). Testing hypotheses about other people: The use of historical knowledge. *Journal of Experimental Social Psychology, 15*, 330-342.

Snyder, M., & Gangestad, S. (1981). Hypothesis-testing processes. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (Vol. 3). Hillsdale, NJ: Erlbaum.

Snyder, M., & Swann, W. B., Jr. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology, 36*, 1202-1212.

Srull, T. K., & Wyer, R. S., Jr. (1986). The role of chronic and temporary goals in social information processing. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 503-549). New York: Guilford Press.

Trope, Y., & Bassok, M. (1982). Confirmatory and diagnosing strategies in social information gathering. *Journal of Personality and Social Psychology, 43*, 22-34.

Trope, Y., & Bassok, M. (1983). Information-gathering strategies in hypothesis-testing. *Journal of Experimental Social Psychology, 19*, 560-576.

Trope, Y., Bassok, M., & Alon, E. (1984). The questions lay interviewers ask. *Journal of Personality, 52*, 90-106.

Trope, Y., & Mackie, D. M. (1987). Sensitivity to alternatives in social hypothesis-testing. *Journal of Experimental Social Psychology, 23*, 445-459.

Wilson, P. (1981). Taking the library out of library education. *American Librarian, 12*, 321-325.

Eugene Borgida is Professor of Psychology and Adjunct Professor of Law and Political Science at the University of Minnesota. His primary research areas are social cognition and psychology and law.

Kenneth G. DeBono is Assistant Professor of Psychology at Union College. His research interests include functional theories of attitudes and individual differences in social cognitive processes.