

Hypothesis-Testing Processes in Social Interaction

Mark Snyder and William B. Swann, Jr.
University of Minnesota

This research is concerned with the processes by which individuals use social interaction to actively test hypotheses about other people. In four separate empirical investigations, female participants were provided with hypotheses about the personal attributes of other individuals ("targets"). Participants then prepared to test these hypotheses (i.e., that their targets were extraverts or that their targets were introverts) by choosing a series of questions to ask their targets in a forthcoming interview. In each investigation, participants planned to test these hypotheses by preferentially searching for behavioral evidence that would confirm these hypotheses. Moreover, these search procedures channeled social interaction between participants and targets in ways that caused the targets to provide actual behavioral confirmation for the participants' hypotheses. A theoretical analysis of the psychological processes believed to underlie and generate both the preferential search for hypothesis-confirming behavioral evidence and the interpersonal consequences of hypothesis-testing activities is presented.

In the course of social relationships, individuals often attempt to make judgments about the personal attributes of other people. At times, this quest for knowledge may involve the testing of hypotheses about other people. When we form our early impressions of new acquaintances, we may wish to test hypotheses based upon our expectations about their personal dispositions (Is this new acquaintance as friendly as a mutual friend has

led me to believe? Is that new acquaintance as boring as every other graduate of the same college?). When we question the accuracy of existing beliefs about friends, we may wish to test hypotheses based on alternate interpretations of their natures (Is this friend whom I have always liked really as mean-tempered as everyone now tells me? Is that friend's unexpected change in behavior a sign of a corresponding change in character?). In these and other circumstances in which we form hypotheses about other people, we may use our subsequent social interactions as opportunities to collect behavioral evidence with which to test these hypotheses.

Having formed a hypothesis about another person, how might an individual use social interaction to actually test that hypothesis? Consider the case of an individual who wishes to test the hypothesis that another person is friendly and sociable. In conversation, the individual might ask that person a series of questions designed to determine whether or not that person's actual behavior and life experiences match those of a characteristically sociable and outgoing person. In choosing these questions, the individual may adopt one

This research and the preparation of this manuscript were supported in part by National Science Foundation Grants SOC 75-13872, "Cognition and Behavior: When Belief Creates Reality," and BNS 77-11346, "From Belief to Reality: Cognitive, Behavioral, and Interpersonal Consequences of Social Perception," to Mark Snyder. We thank Steve Anderson, Bruce Campbell, Bill Gavin, Pat Hanafan, Steve Kramasz, John Lardy, and LaRaye Osborne, who assisted in the empirical phases of these investigations. We thank Berna Skrypnik, Seymour Urano-witz, and D. Kendzierski for their careful reading of this manuscript.

William B. Swann, Jr., is now at the University of Texas at Austin.

Requests for reprints should be sent to Mark Snyder, Laboratory for Research in Social Relations, Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, Minnesota 55455.

of at least three hypothesis-testing strategies. In one strategy, the individual might preferentially search for behavioral evidence that would tend to confirm the hypothesis under scrutiny. Thus, the individual might devote most of the conversation to probing for instances of sociable and outgoing behavior. For example, the individual might ask about those times when the person went to parties, those times when the person sought out new friends, and so forth. In another strategy, the individual might preferentially search for evidence that would disconfirm the hypothesis. Thus, the individual might devote most of the conversation to probing for instances of shy and retiring behavior. For example, the individual might ask about those times when the person wanted to spend time alone or those times when the person avoided meeting new people. In a third strategy, the individual might search for hypothesis-confirming and hypothesis-disconfirming evidence with equal diligence. Thus, the individual might devote equal amounts of the conversation to probing for instances of friendly-sociable and shy-retiring behaviors.

What strategies do individuals actually formulate to test hypotheses about other individuals with whom they interact? Do individuals systematically adopt "confirmatory" strategies and preferentially search for evidence that would confirm their beliefs? Or, do individuals systematically adopt "disconfirmatory" strategies and preferentially search for evidence that would disconfirm their beliefs? Or, do individuals adopt "equal opportunity" strategies and search for confirming and disconfirming behavioral evidence with equal diligence? We have sought answers to these questions in our empirical investigations of hypothesis-testing processes.

Investigation 1:

Formulation of Strategies for Hypothesis Testing

The initial investigation examined strategies that individuals formulate to test hypotheses about others with whom they anticipate social interaction. Each participant received a hypothesis about another individual (the target)

and then prepared to test the hypothesis by planning a series of questions to ask the target. Some participants attempted to assess how extraverted the target was; other participants attempted to assess how introverted the target was. Each participant also received information designed to influence estimates of the likelihood that the hypothesis would prove to be accurate or inaccurate. We included this factor because hypotheses vary in the likelihood that they will prove true; accordingly, we wanted to assess the impact of this factor on hypothesis-testing strategies.

Method

Participants

Participants in these investigations were female undergraduates at the University of Minnesota who received extra credit in their introductory psychology course. The use of participants of one sex was essentially one of convenience; when this research was initiated, there was a substantially greater availability of female participants than male participants. Fifty-eight women participated individually in the first investigation.

Procedure

The experimenter informed participants that they were in an investigation of how people come to understand each other. The experimenter explained that one way to learn about other people is to ask them questions about their likes and dislikes, their favorite activities, their life experiences, and their feelings about themselves. Each participant would attempt to find out about another person (supposedly waiting in another room) by asking questions designed to determine whether that person was the type whose personality was outlined on a card provided by the experimenter. These personality profiles provided the participants with hypotheses about the other individual.

The hypotheses. Participants randomly assigned to the *extravert hypothesis* conditions were instructed to assess the extent to which the target's behavior and experiences matched those of a prototypic extravert. According to the personality profile:

Extraverts are typically outgoing, sociable, energetic, confident, talkative, and enthusiastic. Generally confident and relaxed in social situations, this type of person rarely has trouble making conversation with others. This type of person makes friends quickly and easily and is usually able to make a favorable impression on others. This type of person is usually seen by others as characteristically warm and friendly.

Participants randomly assigned to the *introvert hypothesis* conditions were instructed to determine the extent to which the target's behavior and experiences matched those of a prototypic introvert. According to the personality profile:

Introverts are typically shy, timid, reserved, quiet, distant, and retiring. Usually this type of person would prefer to be alone reading a book or have a long serious discussion with a close friend rather than to go to a loud party or other large social gathering. Often this type of person seems awkward and ill at ease in social situations, and consequently is not adept in making good first impressions. This type of person is usually seen by others as characteristically cool and aloof.

Certainty of the hypothesis. The experimenter also provided participants with information about the (supposed) origins of the profile. Participants assigned randomly to the *high certainty* conditions learned that "the personality profile is a summary of the results of a personality test the other person took last week. Thus, according to the test results, the person whom you will interview is an extravert/introvert." The intent here was to give the hypothesis some credibility by having it "emerge" from the target's own actions. Moreover, pretesting had indicated that undergraduates at the University of Minnesota have considerable faith in the validity of personality assessment procedures.

The intent in the *low certainty* conditions was to make clear that the hypothesis had no connection to any actions of the target. These participants learned that "the personality profile is a description of a type of person familiar to us all—the extravert/introvert. You are to find out how well this profile describes the person you interview." They were given no reasons to believe that the hypothesis was either true or false. Their task simply was to discover whether their target was the type of person described in the personality profile.

Formulating a hypothesis-testing strategy. The experimenter then explained that the profile (the hypothesis) dealt in abstract generalities and global characteristics. However, getting to know someone involves finding out concrete information and specific facts about what that person actually thinks, feels, and does. Accordingly, the participant would choose 12 questions that would help find out whether the target's specific beliefs, attitudes, and actions in life situations matched the general characteristics described in the profile.

The experimenter then provided participants with a list of 26 "Topic Areas Often Covered by Interviewers" from which to choose their 12 questions. The questions on the topic sheet inquired about a wide range of beliefs, feelings, and actions within the domains of personal experience and interpersonal relationships. Nine undergraduate rater-judges had previously classified these questions into three categories.

1. *Extraverted questions.* These 11 questions were

ones that the majority of the rater-judges thought would typically be asked of people *already known* to be extraverts, for example, "What would you do if you wanted to liven things up at a party?" "What kind of situations do you seek out if you want to meet new people?" "In what situations are you most talkative? What is it about these situations that makes you like to talk?"

2. *Introverted questions.* According to the majority of the rater-judges, these 10 questions would characteristically be asked of individuals *already known* to be introverts, for example, "In what situations do you wish you could be more outgoing?" "What factors make it hard for you to really open up to people?" "What things do you dislike about loud parties?"

3. *Neutral questions.* The 5 questions for which there was no consensus that they were extraverted questions or introverted questions and those classified by the majority of the rater-judges as irrelevant to introversion and extraversion were classified as neutral questions, for example, "What kinds of charities do you like to contribute to?" "What are your career goals?" "What do you think the good and bad points of acting friendly and open are?"

Participants selected the 12 questions that they estimated would provide them with the information to best test the hypothesis about the target. The experimenter then informed each participant that the interview would not actually take place and thoroughly debriefed each participant.

Results

What strategies did participants formulate to test hypotheses about targets with whom they anticipated social interaction? Did they plan to preferentially search for evidence that would confirm the hypothesis? Did they plan to preferentially search for evidence that would disconfirm the hypothesis? Or were there no systematic preferences in the participants' strategies?

To answer these questions, we examined the numbers of extraverted, introverted, and neutral questions (for means, see Table 1) that participants planned to ask their targets. A 2 (extravert hypothesis – introvert hypothesis) \times 2 (high certainty – low certainty) multivariate analysis of variance (Overall & Klett, 1972) yielded a highly reliable main effect of the manipulation of the participants' hypotheses, multivariate $F(3, 52) = 11.5, p < .00001$. However, choice of questions was not noticeably affected by the certainty of the hypothesis, multivariate $F(3, 52) < 1$, nor was there any interaction between the hy-

pothesis and certainty factors, multivariate $F(3, 52) = 1.63, ns$.

To specify the nature of the main effect of the participants' hypotheses on their choices of questions, we examined the outcomes of univariate analyses of variance. Participants planned to ask extraverted questions more frequently when preparing to test the hypothesis that their targets were extraverted individuals than when preparing to test the hypothesis that their targets were introverted individuals, $F(1, 54) = 33.04, p < .001$. Similarly, participants chose to ask introverted questions more frequently when planning to test the hypothesis that their targets were introverts than when preparing to test the hypothesis that targets were extraverts, $F(1, 54) = 33.75, p < .001$. Finally, participants chose neutral questions with equal frequency whether they were testing the extravert or the introvert hypothesis, $F(1, 54) = 1.44, ns$.

Discussion

The initial investigation provided evidence that individuals will systematically formulate confirmatory strategies for testing hypotheses about other people. To test the hypothesis that their targets were extraverts, participants were particularly likely to choose to ask those questions that one typically asks of people already known to be extraverts. Similarly, to test the hypothesis that their targets were introverts, participants were particularly likely to choose to ask those questions that one typically asks of people already known to be introverts. Moreover, participants were as likely to plan to search preferentially for evidence that would confirm the hypothesis when they had no reason to believe that the hypothesis was true as when they had some reason to anticipate that the hypothesis was true.

Investigation 2:

Consequences of Confirmatory Hypothesis Testing

Of what consequence is the preferential search for confirming evidence? What would happen if we allowed participants to actually interview their targets and "collect the data"

Table 1
Formulating Strategies for Hypothesis Testing

Type of question	Investigation 1				Investigation 2		Investigations 1 & 2 comparison means		Investigation 3		Investigation 4	
	Extravert hypothesis		Introvert hypothesis		Extravert hypothesis	Introvert hypothesis	Extravert hypothesis	Introvert hypothesis	Many extraverts	Few extraverts	\$25 extravert hypothesis	\$25 introvert hypothesis
	High certainty (n = 14)	Low certainty (n = 15)	High certainty (n = 14)	Low certainty (n = 15)	(n = 20 dyads)	(n = 20 dyads)	(n = 15)	(n = 15)	(n = 15)	(n = 15)	(n = 15)	(n = 15)
Extraverted ^a	6.93	7.67	4.64	4.33	7.25	4.80	7.43	4.60	6.93	6.07	7.20	5.07
<i>M</i>	2.05	1.72	1.69	1.95	2.24	1.73	2.01	1.81	1.33	1.28	2.14	1.67
Introverted ^b	2.43	2.47	5.57	5.60	2.80	5.90	2.66	5.77	3.00	3.87	2.93	4.93
<i>M</i>	2.17	1.99	1.87	2.16	2.31	1.62	2.15	1.85	1.77	1.30	2.25	1.87
Neutral ^c	2.64	1.87	1.78	2.07	1.95	1.30	1.91	1.63	2.07	2.07	1.87	2.00
<i>M</i>	.84	1.18	.80	1.03	.89	1.08	1.01	1.11	.80	1.03	.91	.92
<i>SD</i>												

^a Range = 0-11. Higher means indicate greater numbers of extraverted questions (as identified by the rater-judges) chosen for purposes of hypothesis testing.
^b Range = 0-10. Higher means indicate greater numbers of introverted questions (as identified by the rater-judges) chosen for purposes of hypothesis testing.
^c Range = 0-5. Higher means indicate greater numbers of neutral questions (as identified by the rater-judges) chosen for purposes of hypothesis testing.

that their hypothesis-testing activities would provide them? Would these evidence-gathering procedures generate behaviors that would confirm the hypotheses? Would targets who are being "tested" for extraversion actually come to behave in relatively sociable and outgoing fashion and targets who are being "tested" for introversion actually come to behave in relatively shy and reserved fashion? After all, the more often one inquires about another's person's extraversion, the more opportunities that person will have to provide instances of extraverted behaviors. Similarly, the more often one inquires about another person's introversion, the more often that person will have opportunities to provide instances of introverted behaviors. That is, a confirmatory hypothesis-testing strategy may constrain interaction in ways that cause the target to provide actual *behavioral confirmation* for the hypothesis being tested. In the second experimental investigation, participants first formulated their hypothesis-testing strategies and then carried out these strategies by actually interviewing their targets.

Method

Participants

Participants ($N = 80$) were scheduled in pairs of previously unacquainted individuals who were instructed to arrive at separate experimental rooms located on different corridors. Each participant was assigned randomly to one of two "roles," interviewer or target.

Procedure

The procedure was identical to that of the initial investigation except that participants assigned to the interviewer role actually interviewed participants assigned to the target role. Half of the interviewer-participants ($n = 20$) were instructed to assess the extent to which their target's behavior and experiences matched those of a prototypic extravert. The other half of the interviewer-participants ($n = 20$) were instructed to assess the extent to which their target's behavior and experiences matched those of a prototypic introvert. The instructions to interviewer-participants in the extravert hypothesis and introvert hypothesis conditions were identical to those of the initial investigation. As in the low certainty conditions of the initial investigation, interviewer-participants were given no reasons to anticipate that the

hypothesis would prove accurate or inaccurate. Interviewer-participants were simply instructed to assess the extent to which their targets were like a familiar, but hypothetical, type of person.

While interviewer-participants were choosing their questions from the topic sheet, the experimenter informed participants assigned to the role of target that they would be interviewed by another student. Targets were simply instructed to answer all the questions in as informative, open, and candid a manner as possible.

Each dyad then participated in an interview in which the interviewer-participant asked 12 questions and the target-participant answered these 12 questions. Interviews were conducted by means of microphones and headphones connected through a Sony TC-570 stereophonic tape recorder. Each participant's voice was recorded on a separate channel of the tape. The experimenter then thoroughly debriefed the participants.

To assess the extent to which the answers of the target-participants provided behavioral confirmation for the attributes of the interviewer-participants' hypotheses, six male and six female judges listened to tape recordings of the interviews. These listener-judges were unaware of the purpose of this investigation and knew nothing of the hypotheses being tested by the interviewer-participants. They heard only the track of the tape containing the target-participants' voices. Specifically, they listened to two segments of each interview: the target-participant's answers to three questions from the beginning of the interview and the target-participant's answers to three questions from the end of the interview. Listener-judges then rated each target-participant on 10 6-point bipolar scales: talkative-quiet; unsociable-sociable; friendly-unfriendly; poised-awkward; introverted-extraverted; enthusiastic-apatetic; shy-outgoing; energetic-relaxed; cold-warm; unconfident-confident. These 10 attributes were chosen from the profiles of the prototypic extravert and prototypic introvert that constituted the interviewer-participants' hypotheses. Accordingly, these dependent measures provide indexes of the extent to which specific attributes that define the interviewer-participants' hypotheses were actually reflected in the target-participants' behavior (as perceived by the listener-judges) in this interview context.

Results

We examined the effects of the manipulation of the interviewer-participants' hypotheses on (a) the hypothesis-testing strategies formulated by the interviewer-participants and (b) the target-participants' behavioral self-presentation during the interviews, as measured by the listener-judges' evaluations of the tape recordings.

Interviewer-Participants' Hypothesis-Testing Strategies

Interviewer-participants appear to have formulated confirmatory hypothesis-testing strategies. A multivariate analysis of variance revealed a reliable main effect of the manipulation of the participants' hypotheses, multivariate $F(3, 36) = 7.92, p < .0005$. Univariate analyses of variance on the individual dependent measures (for means, see Table 1) specified the nature of this outcome. Interviewer-participants chose to ask extraverted questions more frequently when planning to test the extravert hypothesis than when planning to test the introvert hypothesis, $F(1, 38) = 14.91, p < .001$. Interviewer-participants chose to ask introverted questions more frequently when preparing to test the introvert hypothesis than when planning to test the extravert hypothesis, $F(1, 38) = 28.18, p < .001$. In addition, interviewer-participants in the extravert hypothesis condition chose neutral questions more frequently than did their counterparts in the introvert hypothesis condition, $F(1, 38) = 4.32, p < .05$.

Target-Participants' Behavioral Confirmation

Interviewer-participants tested hypotheses about their targets by preferentially searching for evidence that would confirm these hypotheses. Moreover, during the interview itself, target-participants came to behave in ways that appeared to confirm specific attributes of the hypotheses being tested by the interviewer-participants. A multivariate analysis of variance (with the 10 dimensions used by the listener-judges as multiple correlated dependent measures) revealed that the listener-judges did view the target-participants in the extravert hypothesis condition quite differently from target-participants in the introvert hypothesis condition, multivariate $F(10, 29) = 2.37, p = .034$.

The nature of the differences detected by the listener-judges may be inferred from group differences on the individual dependent measures. Univariate analyses of variance revealed group differences reliable at better than the .05 level for four of the attributes that had defined the interviewer-participants' hypoth-

eses. Target-participants in the extravert hypothesis condition were regarded as more extraverted, $F(1, 38) = 4.70, p = .036$; confident, $F(1, 38) = 8.41, p = .006$; poised, $F(1, 38) = 12.78, p = .001$; and energetic, $F(1, 38) = 6.59, p = .014$, than target-participants in the introvert hypothesis condition. Two other attributes, when considered individually, yielded group differences reliable at better than the .10 level. Target-participants in the extravert hypothesis condition were seen as more outgoing, $F(1, 38) = 3.55, p = .067$, and more enthusiastic, $F(1, 38) = 3.12, p = .085$, than those in the introvert hypothesis condition. Group differences for the remaining four attributes are all in the same direction. Target-participants in the extravert hypothesis condition were rated as more (although not reliably more) sociable, warm, friendly, talkative, all $F_s \leq 2.36, p_s \geq .13$. And, if one simply sums the 10 dimensions to create an overall measure (with an internal consistency of .95, as assessed by coefficient alpha), a univariate analysis of variance reveals that target-participants in the extravert hypothesis condition presented themselves in more extraverted fashion during the interviews than did participants in the introvert hypothesis condition, $F(1, 38) = 4.56, p = .04$.

Evidently, answers of the target-participants to the interviewer-participants' questions did provide behavioral confirmation for the hypotheses being tested by the interviewer-participants. Moreover, it should be recalled that these behavioral differences were detectable to naive listener-judges who had access only to tape recordings of just the target-participants contributions to the interviews.

Discussion

In this investigation, we were able to witness these stages of the process of hypothesis testing in social interaction: the interviewer-participants' formulation of confirmatory strategies, the interviewer-participants' use of these search procedures in their interviews, and the target-participants' behavioral confirmation of the interviewer-participants' hypotheses. But, did the interviewer-participants regard the hypotheses as having been con-

firmed by the actions of the target-participants? Although this investigation does not answer this question directly, other research (e.g., Swann, 1978) has demonstrated that after interacting with other people for purposes of testing hypotheses, individuals do regard their hypotheses as having been confirmed.

It appears that the critical link in the chain of events of hypothesis testing in social interaction is the formulation of a confirmatory strategy. Once the individual chooses to search preferentially for hypothesis-confirming evidence, the actual behavior of the target then may be constrained in ways that actually provide hypothesis-confirming evidence. Accordingly, we next focused on defining the boundary conditions within which individuals formulate confirmatory hypothesis-testing strategies.

Investigations 3 and 4:

In Search of the Limits of Confirmatory Hypothesis Testing

In the initial investigation, it mattered not at all whether participants had any reason to suspect that the hypothesis would prove to be accurate. When participants had no reason to believe that the hypothesis would prove accurate, they were as likely to plan to search preferentially for confirming evidence as when they had some reason to suspect that the hypothesis might describe the target accurately. How pervasive is this hypothesis-testing strategy? In the third and fourth investigations, we attempted to identify circumstances in which individuals would avoid confirmatory strategies.

Method: Investigation 3

Will individuals avoid confirmatory hypothesis-testing strategies if they have compelling reasons to believe that the hypothesis may prove to be inaccurate? Participants in the third investigation ($N = 30$) chose 12 questions to test the hypothesis that their targets were extraverts. The procedure was identical to that of the initial investigation, except that participants received concrete information designed to graphically convey the likelihood that the hypothesis would prove accurate or inaccurate.

To make it seem unlikely that the targets were

extraverts, we informed participants in the few extraverts condition ($n = 15$) that:

She is a sophomore, a psychology major, and a member of a sorority. From a recent study we did, we know that very few of the 30 members of this particular sorority are extraverts. That is, of the 30 girls in the sorority, 7 of them are extraverts. Your task is to find out if she is one of the very few extraverts in her sorority.

To make it seem likely that the targets were extraverts, we informed participants in the many extraverts condition ($n = 15$) that:

She is a sophomore, a psychology major, and a member of a sorority. From a recent study we did, we know that most of the 30 members of this particular sorority are extraverts. That is, of the 30 girls in the sorority, 23 of them are extraverts. Your task is to find out if she is one of the many extraverts in her sorority.

To assess understanding of the implications of the composition of the sorority, we had participants answer, on a 6-point scale, the question, "In view of the number of extraverts in her sorority, how likely is it that the person you are about to interview is an extravert?" We included this question because we were aware that there are times when people fail to appreciate the implications of statistical information about other people (e.g., Nisbett & Borgida, 1975).

Results: Investigation 3

Did it matter to participants whether their targets were one of very many or very few extraverts in the sorority? They appear to have understood the implications of the information about the sorority. Participants in the many extraverts condition estimated that it was much more likely that their targets were extraverts ($M = 5.0$) than did participants in the few extraverts condition ($M = 3.8$), $F(1, 28) = 12.39$, $p < .001$. But, did this information affect the formulation of hypothesis-testing strategies?

As standards of comparison for assessing the extent to which participants in the many extraverts and the few extraverts conditions formulated confirmatory strategies, we created extravert hypothesis and introvert hypothesis comparison conditions. The extravert hypothesis comparison condition ($n = 35$) was created by combining participants in the extravert hypothesis—low certainty condition of the first investigation and participants in the

extravert hypothesis condition of the second investigation. The introvert hypothesis comparison condition ($n = 35$) was created by combining the introvert hypothesis - low certainty condition of the first investigation and the introvert hypothesis condition of the second investigation. These comparison conditions provide our best estimates of the mean numbers of extraverted, introverted, and neutral questions chosen to test hypotheses when there are no reasons to expect that the hypotheses will prove accurate or inaccurate.¹

How did the hypothesis-testing strategies formulated by participants in the many extraverts and few extraverts conditions compare with those of participants in the extravert hypothesis and introvert hypothesis comparison conditions? A one-way multivariate analysis of variance revealed reliable differences in the pattern of extraverted, introverted, and neutral questions chosen by the four groups, multivariate $F(9, 229) = 5.18, p < .001$. Subsequent univariate analyses of variance revealed reliable between-group differences for mean numbers of extraverted questions and introverted questions chosen, $F_s(3, 96) = 16.13$ and 17.62 , respectively, $p_s < .001$, but no reliable between-groups differences in mean numbers of neutral questions chosen, $F(3, 96) = 1.03, ns$.

To infer the nature of these between-groups differences, we then examined the pattern of the group means for extraverted and introverted questions. It appears that participants planned to ask equally many *extraverted questions* (for means, see Table 1) whether they were in the many extraverts, few extraverts, or extraverts hypothesis comparison conditions; and participants in the many extraverts, few extraverts, and extravert hypothesis comparison conditions all planned to ask more extraverted questions than participants in the introvert hypothesis comparison condition, $F(1, 96) = 42.10, p < .001$.² Moreover, participants planned to ask equally few *introverted questions* whether they were in the many extraverts, few extraverts, or extravert hypothesis comparison conditions; and participants in the many extraverts, few extraverts, and extravert hypothesis comparison conditions all planned to ask fewer introverted

questions than participants in the introvert hypothesis comparison condition, $F(1, 96) = 48.54, p < .001$.³

Method: Investigation 4

Would participants formulate confirmatory hypothesis-testing strategies if offered substantial incentives to test hypotheses as accurately as possible? We made the following offer to participants in the fourth investigation ($N = 30$):

To give you a little incentive, we are offering \$25 to the person who develops the set of questions that will tell the most about the extraversion/introversion of the interviewee. You should therefore try to be as accurate as possible in finding out what the interviewee is like.

Participants in the \$25 extravert hypothesis condition ($n = 15$) and those in the \$25 introvert hypothesis condition ($n = 15$) then chose 12 questions to test

¹ Before combining results from the relevant conditions of the first and the second investigations, we first ascertained that our "hypothesis-testing" effect (that is, preferential choice of hypothesis-confirming questions) did not differ across the two investigations. Indeed, if we perform a 2 (introvert hypothesis - extravert hypothesis) \times 2 (Investigation 1 - Investigation 2) multivariate analysis of variance with numbers of extraverted, introverted, and neutral questions as dependent variables, there are neither main nor interaction effects involving the investigation factor, multivariate $F_s(3, 64) \leq 1.06, ns$; there is, of course, a reliable main effect of the hypothesis factor, multivariate $F(3, 64) = 13.71, p < .001$.

² This F value is the outcome of a single contrast with the following weights: many extraverts = +1; few extraverts = +1; extravert hypothesis = +1; introvert hypothesis = -3. Not only is this contrast highly significant but it accounts for 87.01% of the systematic between-conditions variance. Moreover, the many extraverts, few extraverts, and extravert hypothesis comparison means for extraverted questions do not differ from each other at the .05 level of confidence, using Scheffé's procedure for multiple comparisons.

³ This F value is the outcome of a single contrast with the following weights: many extraverts = -1; few extraverts = -1; extravert hypothesis = -1; introvert hypothesis = +3. Not only is this contrast highly reliable but it accounts for 91.84% of the systematic between-conditions variance. Moreover, the many extraverts, few extraverts, and extravert hypothesis comparison means for introverted questions do not differ from each other at the .05 level of confidence, using Scheffé's procedure for multiple comparisons.

the hypotheses that their targets were extraverts or introverts, respectively.⁴

Results: Investigation 4

Did participants avoid the preferential search for confirming evidence when offered a \$25 incentive for accuracy? Or, more modestly, were these participants any less likely to formulate confirmatory strategies than individuals who had not been offered such large incentives for accuracy (i.e., the extravert hypothesis and introvert hypothesis comparison conditions)? Evidently not. When we entered choices of extraverted, introverted, and neutral questions into a 2 (extravert hypothesis – introvert hypothesis) \times 2 (\$25–comparison) multivariate analysis of variance, there emerged a reliable main effect of the hypothesis manipulation, multivariate $F(3, 94) = 16.36, p < .0001$, but neither a reliable main effect of the investigation factor nor a reliable interaction between the hypothesis and investigation factors, multivariate $F_s < 1$.

Moreover, separate univariate analyses (for means, see Table 1) indicated that participants in the \$25 extravert hypothesis condition planned to ask more extraverted questions than participants in the \$25 introvert hypothesis condition, $F(1, 28) = 9.25, p = .005$; and participants in the \$25 introvert hypothesis condition planned to ask more introverted questions than participants in the \$25 extravert hypothesis condition, $F(1, 28) = 7.01, p = .013$. Apparently, the offer of substantial monetary incentives was not sufficient even to diminish, let alone override, the propensity to search preferentially for confirming evidence.

Discussion: Investigations 3 and 4

In each investigation, we have observed the formulation of confirmatory hypothesis-testing strategies. It seemed not to matter to participants where their hypotheses originated (Investigation 1), how likely it was that their hypotheses would prove accurate (Investigation 3), or whether substantial incentives for accurate hypothesis testing were offered (Investigation 4). In each case, participants planned to preferentially search for evidence that would tend to confirm their hypotheses.

It is true that all participants in these investigations were females. However, we know of no reason why the outcomes ought not be generalizable to males. In fact, other research on hypothesis testing in social interaction has documented the preferential search for hypothesis-confirming evidence with male participants (Swann, 1978). Nonetheless, we recognize that we are in no position to claim that individuals always adopt confirmatory hypothesis-testing strategies. At the very least, we can assert that we have yet to identify any procedure that will induce individuals to eschew such strategies in favor of either disconfirmatory or equal opportunity strategies.

Hypothesis Testing: A Theoretical Analysis

Why did individuals who participated in these empirical investigations construct and enact confirmatory strategies for testing hypotheses about other people? We believe that the structure and process of human thought fosters and promotes the ready and willing adoption of confirmatory strategies for hypothesis testing. First of all, there is every reason to believe that it is easier for the individual to think of the target behaving in accord with his or her hypothesized nature than it is for the individual to think of the target violating the hypothesis. Considerable research on concept formation and concept utilization indicates that people prefer and use positive instances of concepts over negative ones (e.g., Hovland & Weiss, 1953). Moreover, confirming instances generally have more impact on inductive conclusions than do disconfirming instances (e.g., Gollub, Rossman, & Abelson, 1973), and covariation between positive instances leads to estimates of greater relationships than does covariation between negative or mixed instances (e.g., Jenkins & Ward, 1965; Smedslund, 1963). Furthermore, in judgments of similarity, individuals preferen-

⁴ We did award \$25 to one participant in the fourth investigation. Our criterion for identifying "accuracy" was somewhat arbitrary. We awarded \$25 to one of the only two participants (she was identified by the flip of a coin) who planned to ask precisely equal numbers of extraverted and introverted questions of their targets.

tially look for common features rather than distinctive features (e.g., Tversky, 1977). Similarly, investigations of logical reasoning (e.g., Wason & Johnson-Laird, 1972) make clear that when attempting to decide whether general propositions (e.g., all Norwegian men are handsome) are true, individuals almost always look for instances that could verify the proposition (e.g., Norwegian men who are handsome) and almost never look for falsifying instances (e.g., nonhandsome men who are Norwegian). Even researchers in the behavioral sciences tend to design empirical investigations that seek to confirm, rather than disconfirm, their hypotheses (Greenwald, 1975). Accordingly, when anticipating what events are to appear as their interaction unfolds, the individual may find it easier to construct mental scenarios of the events of the forthcoming interaction in which the target acts in accord with the individual's hypothesis than to construct scenarios in which the target betrays the hypothesis.

If representations of the target behaviorally confirming the individual's hypothesis are more cognitively "available" than representations of the target violating the hypothesis, then there is every reason to believe that the individual will overestimate the likelihood that the target will, in fact, behave in ways that confirm the hypothesis. Considerable evidence suggests that individuals use "availability" as a heuristic for estimating frequency: Events that are easy to bring to mind are thought to occur with greater frequency than events that are difficult to bring to mind (e.g., Tversky & Kahneman, 1973). If so, by virtue of contemplating the forthcoming interaction with the target in the light of the hypothesis, the individual not only will find it easier to think of the target confirming the hypothesis but also will believe that these hypothesis-confirming actions will occur in great numbers and that these hypothesis-confirming behaviors will be representative of the target's true personal nature.

To the extent that the individual believes that hypothesis-confirming behaviors are typical of the target's activities, he or she may consider it not unreasonable to confine the conversation to those topics about which the

target can provide the most informative and meaningful facts about his or her life. Accordingly, the individual may use their social interaction as an opportunity to collect preferentially evidence that confirms the hypothesis under scrutiny.

Such a preferential evidence-gathering procedure may generate a sample of evidence in which hypothesis-confirming evidence will be overrepresented and hypothesis-disconfirming evidence will be underrepresented. For there is every reason to believe that most people, as targets, will be "generous" in providing specific instances of hypothesis-confirming actions. There is sufficient situation-to-situation variability in human social behavior that most people about whom hypotheses are tested will have behaved, in some situations and at some times, consistently with the hypothesis under scrutiny (e.g., Mischel, 1968). However, these same people probably will have behaved, in other situations and at other times, in ways that would tend to disconfirm that same hypothesis. Accordingly, to the extent that the individual preferentially solicits hypothesis-confirming instances of the target's behavior, such a search procedure will be particularly successful in generating a sample of data in which confirming evidence is overrepresented and in which disconfirming evidence is underrepresented. Of course, it will be this sample of data upon which the individual will base the decision to accept or reject the hypothesis in question. Accordingly, the individual may accept this hypothesis more readily than the "data" of actual events in the target's life truly warrant.

Whatever the ultimate fate of this admittedly speculative analysis of the processes that underlie and generate the preferential search for hypothesis-confirming evidence, the possible social and interpersonal consequences of confirmatory hypothesis-testing strategies cannot be ignored. To the extent that individuals chronically formulate and enact confirmatory strategies for assessing the accuracy of their hypotheses and beliefs about other people, they may create for themselves a world in which hypotheses become self-confirming hypotheses and beliefs become self-perpetuating beliefs (for other demonstrations of the self-

perpetuating nature of beliefs, see Snyder & Swann, 1978; Snyder, Tanke, & Berscheid, 1977; Snyder & Uranowitz, 1978). From this perspective, it becomes easier to understand why so many popular beliefs about other people (in particular, clearly erroneous social and cultural stereotypes) are so stubbornly resistant to change. Even if one were to develop sufficient doubt about the accuracy of these beliefs to proceed to test them actively, one nevertheless might be likely to "find" all the evidence one needs to confirm and retain these beliefs. And, in the end, one may be left with the secure (but totally unwarranted) feeling that these beliefs must be correct because they have survived (what may seem to the individual) perfectly appropriate and even rigorous procedures for assessing their accuracy.

References

- Gollob, H. F., Rossman, B. B., & Abelson, R. P. Social inference as a function of the number of instances and consistency of information presented. *Journal of Personality and Social Psychology*, 1973, 27, 19-33.
- Greenwald, A. G. Consequences of prejudice against the null hypothesis. *Psychological Bulletin*, 1975, 82, 1-20.
- Hovland, C. I., & Weiss, W. Transmission of information concerning concepts through positive and negative instances. *Journal of Experimental Psychology*, 1953, 45, 175-182.
- Jenkins, H. M., & Ward, W. C. Judgment of contingency between responses and outcomes. *Psychological Monographs*, 1965, 79(1, Whole No. 594).
- Mischel, W. *Personality and assessment*. New York: Wiley, 1968.
- Nisbett, R. E., & Borgida, E. Attribution and the psychology of prediction. *Journal of Personality and Social Psychology*, 1975, 32, 932-943.
- Overall, J. E., & Klett, C. J. *Applied multivariate analysis*. New York: McGraw-Hill, 1972.
- Smedslund, J. The concept of correlation in adults. *Scandinavian Journal of Psychology*, 1963, 4, 165-173.
- Snyder, M., & Swann, W. B., Jr. Behavioral confirmation in social interaction: From social perception to social reality. *Journal of Experimental Social Psychology*, 1978, 14, 148-162.
- Snyder, M., Tanke, E. D., & Berscheid, E. Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, 1977, 35, 656-666.
- Snyder, M., & Uranowitz, S. W. Reconstructing the past: Some cognitive consequences of person perception. *Journal of Personality and Social Psychology*, 1978, 36, 941-950.
- Swann, W. B., Jr. *The interpersonal nature of self-conceptions*. Unpublished doctoral dissertation, University of Minnesota, 1978.
- Tversky, A. Features of similarity. *Psychological Review*, 1977, 84, 327-352.
- Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 1973, 5, 207-232.
- Wason, P. C., & Johnson-Laird, P. N. *Psychology of reasoning: Structure and content*. London: D. T. Batsford, 1972.

Received May 1, 1978 ■