Confirmation Bias: A Ubiquitous Phenomenon in Many Guises

Raymond S. Nickerson
Tufts University

Confirmation bias, as the term is typically used in the psychological literature, connotes the seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand. The author reviews evidence of such a bias in a variety of guises and gives examples of its operation in several practical contexts. Possible explanations are considered, and the question of its utility or disutility is discussed.

When men wish to construct or support a theory, how they torture facts into their service! (Mackay, 1852/1932, p. 552)

Confirmation bias is perhaps the best known and most widely accepted notion of inferential error to come out of the literature on human reasoning. (Evans, 1989, p. 41)

If one were to attempt to identify a single problematic aspect of human reasoning that deserves attention above all others, the *confirmation bias* would have to be among the candidates for consideration. Many have written about this bias, and it appears to be sufficiently strong and pervasive that one is led to wonder whether the bias, by itself, might account for a significant fraction of the disputes, altercations, and misunderstandings that occur among individuals, groups, and nations.

Confirmation bias has been used in the psychological literature to refer to a variety of phenomena. Here I take the term to represent a generic concept that subsumes several more specific ideas that connote the inappropriate bolstering of hypotheses or beliefs whose truth is in question.

Deliberate Versus Spontaneous Case Building

There is an obvious difference between impartially evaluating evidence in order to come to an unbiased conclusion and building a case to justify a conclusion already drawn. In the first instance one seeks evidence on all sides of a

Correspondence concerning this article should be addressed to Raymond S. Nickerson, Department of Psychology, Paige Hall, Tufts University, Medford, Massachusetts 02155. Electronic mail may be sent to rnickerson@infonet.tufts.edu.

question, evaluates it as objectively as one can, and draws the conclusion that the evidence, in the aggregate, seems to dictate. In the second, one selectively gathers, or gives undue weight to, evidence that supports one's position while neglecting to gather, or discounting, evidence that would tell against it.

There is a perhaps less obvious, but also important, difference between building a case consciously and deliberately and engaging in case-building without being aware of doing so. The first type of case-building is illustrated by what attorneys and debaters do. An attorney's job is to make a case for one or the other side of a legal dispute. The prosecutor tries to marshal evidence to support the contention that a crime has been committed; the defense attorney tries to present evidence that will support the presumption that the defendant is innocent. Neither is committed to an unbiased weighing of all the evidence at hand, but each is motivated to confirm a particular position. Debaters also would be expected to give primary attention to arguments that support the positions they are defending; they might present counterarguments, but would do so only for the purpose of pointing out their weaknesses.

As the term is used in this article and, I believe, generally by psychologists, *confirmation bias* connotes a less explicit, less consciously one-sided case-building process. It refers usually to unwitting selectivity in the acquisition and use of evidence. The line between deliberate selectivity in the use of evidence and unwitting molding of facts to fit hypotheses or beliefs is a difficult one to draw in practice, but the distinction is meaningful conceptually, and confirmation bias has more to do with the latter than with the former. The

assumption that people can and do engage in case-building unwittingly, without intending to treat evidence in a biased way or even being aware of doing so, is fundamental to the concept.

The question of what constitutes confirmation of a hypothesis has been a controversial matter among philosophers and logicians for a long time (Salmon, 1973). The controversy is exemplified by Hempel's (1945) famous argument that the observation of a white shoe is confirmatory for the hypothesis "All ravens are black," which can equally well be expressed in contrapositive form as "All nonblack things are nonravens." Goodman's (1966) claim that evidence that something is green is equally good evidence that it is "grue"—*grue* being defined as green before a specified future date and blue thereafter—also provides an example. A large literature has grown up around these and similar puzzles and paradoxes. Here this controversy is largely ignored. It is sufficiently clear for the purposes of this discussion that, as used in everyday language, *confirmation* connotes evidence that is perceived to *support*—to increase the credibility of—a hypothesis.

I also make a distinction between what might be called *motivated* and *unmotivated* forms of confirmation bias. People may treat evidence in a biased way when they are motivated by the desire to defend beliefs that they wish to maintain. (As already noted, this is not to suggest intentional mistreatment of evidence; one may be selective in seeking or interpreting evidence that pertains to a belief without being deliberately so, or even necessarily being aware of the selectivity.) But people also may proceed in a biased fashion even in the testing of hypotheses or claims in which they have no material stake or obvious personal interest. The former case is easier to understand in commonsense terms than the latter because one can appreciate the tendency to treat evidence selectively when a valued belief is at risk. But it is less apparent why people should be partial in their uses of evidence when they are indifferent to the answer to a question in hand. An adequate account of the confirmation bias must encompass both cases because the existence of each is well documented.

There are, of course, instances of one wishing to disconfirm a particular hypothesis. If, for example, one believes a hypothesis to be untrue, one may seek evidence of that fact or give undue weight to such evidence. But in such cases, the hypothesis in question is someone else's belief. For the individual who seeks to disconfirm such a hypothesis, a confirmation bias would be a bias to confirm the individual's own belief, namely that the hypothesis in question is false.

A Long-Recognized Phenomenon

Motivated confirmation bias has long been believed by philosophers to be an important determinant of thought and behavior. Francis Bacon (1620/1939) had this to say about it, for example:

> The human understanding when it has once adopted an opinion (either as being the received opinion or as being agreeable to itself) draws all things else to support and agree with it. And though there be a greater number and weight of instances to be found on the other side, yet these it either neglects and despises, or else by some distinction sets aside and rejects; in order that by this great and pernicious predetermination the authority of its former conclusions may remain inviolate. . . . And such is the way of all superstitions, whether in astrology, dreams, omens, divine judgments, or the like; wherein men, having a delight in such vanities, mark the events where they are fulfilled, but where they fail, although this happened much oftener, neglect and pass them by. (p. 36)

Bacon noted that philosophy and the sciences do not escape this tendency.

The idea that people are prone to treat evidence in biased ways if the issue in question matters to them is an old one among psychologists also:

> If we have nothing personally at stake in a dispute between people who are strangers to us, we are remarkably intelligent about weighing the evidence and in reaching a rational conclusion. We can be convinced in favor of either of the fighting parties on the basis of good evidence. But let the fight be our own, or let our own friends, relatives, fraternity brothers, be parties to the fight, and we lose our ability to see any other side of the issue than our own. . . . The more urgent the impulse, or the closer it comes to the maintenance of our own selves, the more difficult it becomes to be rational and intelligent. (Thurstone, 1924, p. 101)

The data that I consider in what follows do not challenge either the notion that people generally like to avoid personally disquieting information or the belief that the strength of a bias in the interpretation of evidence increases with the degree to which the evidence relates directly to a dispute in which one has a personal

stake. They are difficult to reconcile, however, with the view that evidence is treated in a totally unbiased way if only one has no personal interest in that to which it pertains.

The following discussion of this widely recognized bias is organized in four major sections. In the first, I review experimental evidence of the operation of a confirmation bias. In the second, I provide examples of the bias at work in practical situations. The third section notes possible theoretical explanations of the bias that various researchers have proposed. The fourth addresses the question of the effects of the confirmation bias and whether it serves any useful purposes.

Experimental Studies

A great deal of empirical evidence supports the idea that the confirmation bias is extensive and strong and that it appears in many guises. The evidence also supports the view that once one has taken a position on an issue, one's primary purpose becomes that of defending or justifying that position. This is to say that regardless of whether one's treatment of evidence was evenhanded before the stand was taken, it can become highly biased afterward.

Hypothesis-Determined Information Seeking and Interpretation

People tend to seek information that they consider supportive of favored hypotheses or existing beliefs and to interpret information in ways that are partial to those hypotheses or beliefs. Conversely, they tend not to seek and perhaps even to avoid information that would be considered counterindicative with respect to those hypotheses or beliefs and supportive of alternative possibilities (Koriat, Lichtenstein, & Fischhoff, 1980).

Beyond seeking information that is supportive of an existing hypothesis or belief, it appears that people often tend to seek only, or primarily, information that will support that hypothesis or belief in a particular way. This qualification is necessary because it is generally found that people seek a specific type of information that they would expect to find, assuming the hypothesis is true. Also, they sometimes appear to give weight to information that is consistent with a hypothesis but not diagnostic with respect

to it. These generalizations are illustrated by several of the following experimental findings.

Restriction of attention to a favored hypothesis. If one entertains only a single possible explanation of some event or phenomenon, one precludes the possibility of interpreting data as supportive of any alternative explanation. Even if one recognizes the possibility of other hypotheses or beliefs, perhaps being aware that other people hold them, but is strongly committed to a particular position, one may fail to consider the relevance of information to the alternative positions and apply it (favorably) only to one's own hypothesis or belief.

Restricting attention to a single hypothesis and failing to give appropriate consideration to alternative hypotheses is, in the Bayesian framework, tantamount to failing to take likelihood ratios into account. The likelihood ratio is the ratio of two conditional probabilities, $p(D|H_1)/p(D|H_j)$, and represents the probability of a particular observation (or datum) if one hypothesis is true relative to the probability of that same observation if the other hypothesis is true. Typically there are several plausible hypotheses to account for a specific observation, so a given hypothesis would have several likelihood ratios. The likelihood ratio for a hypothesis and its complement, $p(D|H)/p(D|\sim H)$, is of special interest, however, because an observation gives one little evidence about the probability of the truth of a hypothesis unless the probability of that observation, given that the hypothesis is true, is either substantially larger or substantially smaller than the probability of that observation, given that the hypothesis is false.

The notion of diagnosticity reflects the importance of considering the probability of an observation conditional on hypotheses other than the favored one. An observation is said to be *diagnostic* with respect to a particular hypothesis to the extent that it is consistent with that hypothesis and not consistent, or not as consistent, with competing hypotheses and in particular with the complementary hypothesis. One would consider an observation diagnostic with respect to a hypothesis and its complement to the degree that the likelihood ratio, $p(D|H)/p(D|\sim H)$, differed from 1. An observation that is consistent with more than one hypothesis is said not to be diagnostic with respect to those hypotheses; when one considers the probability

of an observation conditional on only a single hypothesis, one has no way of determining whether the observation is diagnostic.

Evidence suggests that people often do the equivalent of considering only $p(D|H)$ and failing to take into account the ratio of this and $p(D|{\sim}H)$, despite the fact that considering only one of these probabilities does not provide a legitimate basis for assessing the credibility of H (Beyth-Marom & Fischhoff, 1983; Doherty & Mynatt, 1986; Doherty, Mynatt, Tweney, & Schiavo, 1979; Griffin & Tversky, 1992; Kern & Doherty, 1982; Troutman & Shanteau, 1977). This tendency to focus exclusively on the case in which the hypothesis is assumed to be true is often referred to as a tendency toward *pseudodi-agnosticity* (Doherty & Mynatt, 1986; Doherty et al., 1979; Fischhoff & Beyth-Marom, 1983; Kern & Doherty, 1982). Fischhoff and Beyth-Marom (1983) have argued that much of what has been interpreted as a confirmation bias can be attributed to such a focus and the consequential failure to consider likelihood ratios.

Preferential treatment of evidence supporting existing beliefs. Closely related to the restriction of attention to a favored hypothesis is the tendency to give greater weight to information that is supportive of existing beliefs or opinions than to information that runs counter to them. This does not necessarily mean completely ignoring the counterindicative information but means being less receptive to it than to supportive information—more likely, for example, to seek to discredit it or to explain it away.

Preferential treatment of evidence supporting existing beliefs or opinions is seen in the tendency of people to recall or produce reasons supporting the side they favor—*my-side bias*—on a controversial issue and not to recall or produce reasons supporting the other side (Baron, 1991, 1995; Perkins, Allen, & Hafner, 1983; Perkins, Farady, & Bushey, 1991). It could be either that how well people remember a reason depends on whether it supports their position, or that people hold a position because they can think of more reasons to support it. Participants in the study by Perkins, Farady, and Bushey were capable of generating reasons for holding a view counter to their own when explicitly asked to do so; this finding led Tishman, Jay, and Perkins (1993) to interpret the

failure to do so spontaneously as a motivational problem as distinct from a cognitive limitation.

Baron (1995) found that, when asked to judge the quality of arguments, many people were likely to rate one-sided arguments higher than two-sided arguments, suggesting that the bias is at least partially due to common beliefs about what makes an argument strong. In keeping with this result, participants in a mock jury trial who tended to use evidence selectively to build one view of what happened expressed greater confidence in their decisions than did those who spontaneously tried to weigh both sides of the case (D. Kuhn, Weinstock, & Flaton, 1994).

When children and young adults were given evidence that was inconsistent with a theory they favored, they often "either failed to acknowledge discrepant evidence or attended to it in a selective, distorting manner. Identical evidence was interpreted one way in relation to a favored theory and another way in relation to a theory that was not favored" (D. Kuhn, 1989, p. 677). Some of Kuhn's participants were unable to indicate what evidence would be inconsistent with their theories; some were able to generate alternative theories when asked, but they did not do so spontaneously. When they were asked to recall their theories and the related evidence that had been presented, participants were likely to recall the evidence as being more consistent with the theories than it actually was. The greater perceived consistency was achieved sometimes by inaccurate recall of theory and sometimes by inaccurate recall of evidence.

Looking only or primarily for positive cases. What is considerably more surprising than the fact that people seek and interpret information in ways that increase their confidence in favored hypotheses and established beliefs is the fact that they appear to seek confirmatory information even for hypotheses in whose truth value they have no vested interest. In their pioneering concept-discovery experiments, Bruner, Goodnow, and Austin (1956) found that participants often tested a hypothesized concept by choosing only examples that would be classified as instances of the sought-for concept if the hypothesis were correct. This strategy precludes discovery, in some cases, that an incorrect hypothesis is incorrect. For example, suppose the concept to be discovered is *small circle* and one's hypothesis is *small red circle*. If one tests the hypothesis by selecting only things that are

small, red, and circular, one will never discover that the class defined by the concept includes also small circular things that are yellow or blue. Several investigators (Levine, 1970; Millward & Spoehr, 1973; Taplin, 1975; Tweney et al., 1980; Wason & Johnson-Laird, 1972) subsequently observed the same behavior of participants testing only cases that are members of the hypothesized category.

Some studies demonstrating selective testing behavior of this sort involved a task invented by Wason (1960) in which people were asked to find the rule that was used to generate specified triplets of numbers. The experimenter presented a triplet, and the participant hypothesized the rule that produced it. The participant then tested the hypothesis by suggesting additional triplets and being told, in each case, whether it was consistent with the rule to be discovered. People typically tested hypothesized rules by producing only triplets that were consistent with them. Because in most cases they did not generate any test items that were inconsistent with the hypothesized rule, they precluded themselves from discovering that it was incorrect if the triplets it prescribed constituted a subset of those prescribed by the actual rule. Given the triplet 2–4–6, for example, people were likely to come up with the hypothesis *successive even numbers* and then proceed to test this hypothesis by generating additional sets of successive even numbers. If the 2–4–6 set had actually been produced by the rule *numbers increasing by 2, numbers increasing in size,* or *any three positive numbers,* the strategy of using only sets of successive even numbers would not reveal the incorrectness of the hypothesis because every test item would get a positive response.

The use only of test cases that will yield a positive response if a hypothesis under consideration is correct not only precludes discovering the incorrectness of certain types of hypotheses with a correct hypothesis, this strategy would not yield as strongly confirmatory evidence, logically, as would that of deliberately selecting tests that would show the hypothesis to be wrong, if it is wrong, and failing in the attempt. To the extent that the strategy of looking only for positive cases is motivated by a wish to find confirmatory evidence, it is misguided. The results this endeavor will yield will, at best, be consistent with the hypothesis, but the confirmatory evidence they provide will not be as

compelling as would the failure of a rigorous attempt at disconfirmation.

This point is worth emphasizing because the psychological literature contains many references to the confirmatory feedback a participant gets when testing a hypothesis with a positive case. These references do not generally distinguish between confirmatory in a logical sense and confirmatory in a psychological sense. The results obtained by Wason (1960) and others suggest that feedback that is typically interpreted by participants to be strongly confirmatory often is not logically confirmatory, or at least not strongly so. The "confirmation" the participant receives in this situation is, to some degree, illusory. This same observation applies to other studies mentioned in the remainder of this article.

In an early commentary on the triplet-rule task, Wetherick (1962) argued that the experimental situation did not reveal the participants' intentions in designating any particular triplet as a test case. He noted that any test triplet could either conform or not conform to the rule, as defined by the experimenter, and it could also either conform or not conform to the hypothesis being considered by the participant. Any given test case could relate to the rule and hypothesis in combination in any of four ways: conform–conform, conform–not conform, not conform–conform, and not conform–not conform. Wetherick argued that one could not determine whether an individual was intentionally attempting to eliminate a candidate hypothesis unless one could distinguish between test cases that were selected because they conformed to a hypothesis under consideration and those that were selected because they did not.

Suppose a participant selects the triplet 3–5–7 and is told that it is consistent with the rule (the rule being *any three numbers in ascending order*). The participant might have chosen this triplet because it conforms to the hypothesis being considered, say *numbers increasing by two,* and might have taken the positive response as evidence that the hypothesis is correct. On the other hand, the participant could have selected this triplet in order to eliminate one or more possible hypotheses (e.g., *even numbers ascending; a number, twice the number, three times the number*). In this case, the experimenter's positive response would constitute the discon-

firming evidence (with respect to these hypotheses) the participant sought.

Wetherick (1962) also pointed out that a test triplet may logically rule out possible hypotheses without people being aware of the fact because they never considered those hypotheses. A positive answer to the triplet 3–5–7 logically eliminates *even numbers ascending* and *a number, twice the number, three times the number,* among other possibilities, regardless of whether the participant thought of them. But of course, only if the triplet was selected with the intention of ruling out those options should its selection be taken as an instance of a falsification strategy. Wetherick's point was that without knowing what people have in mind in making the selections they do, one cannot tell whether they are attempting to eliminate candidates from further consideration or not.

Wason (1962, 1968/1977) responded to this objection with further analyses of the data from the original experiment and data from additional experiments showing that although some participants gave evidence of understanding the concept of falsification, many did not. Wason summarized the findings from these experiments this way: "there would appear to be compelling evidence to indicate that even intelligent individuals adhere to their own hypotheses with remarkable tenacity when they can produce confirming evidence for them" (1968/1977, p. 313).

In other experiments in which participants have been asked to determine which of several hypotheses is the correct one to explain some situation or event, they have tended to ask questions for which the correct answer would be *yes* if the hypothesis under consideration were true (Mynatt, Doherty, & Tweney, 1977; Shaklee & Fischhoff, 1982). These experiments are among many that have been taken to reveal not only a disinclination to test a hypothesis by selecting tests that would show it to be false if it is false, but also a preference for questions that will yield a positive answer if the hypothesis is true.

Others have noted the tendency to ask questions for which the answer is *yes* if the hypothesis being tested is correct in the context of experiments on personality perception (Hodgins & Zuckerman, 1993; Schwartz, 1982; Strohmer & Newman, 1983; Trope & Bassock, 1982, 1983; Trope, Bassock, & Alon, 1984; Zuckerman, Knee, Hodgins, & Miyake, 1995). Fischhoff and Beyth-Marom (1983) also noted the possibility that participants in such experiments tend to assume that the hypothesis they are asked to test is true and select questions that would be the least awkward to answer if that is the case. For instance, participants asked assumed extroverts (or introverts) questions that extroverts (or introverts) would find particularly easy to answer.

Overweighting positive confirmatory instances. Studies of social judgment provide evidence that people tend to overweight positive confirmatory evidence or underweight negative discomfirmatory evidence. Pyszczynski and Greenberg (1987) interpreted such evidence as supportive of the view that people generally require less hypothesis-consistent evidence to accept a hypothesis than hypothesis-inconsistent information to reject a hypothesis. These investigators argued, however, that this asymmetry is modulated by such factors as the degree of confidence one has in the hypothesis to begin with and the importance attached to drawing a correct conclusion. Although they saw the need for accuracy as one important determinant of hypothesis-evaluating behavior, they suggested that other motivational factors, such as needs for self-esteem, control, and cognitive consistency, also play significant roles.

People can exploit others' tendency to overweight (psychologically) confirming evidence and underweight disconfirming evidence for many purposes. When the mind reader, for example, describes one's character in more-or-less universally valid terms, individuals who want to believe that their minds are being read will have little difficulty finding substantiating evidence in what the mind reader says if they focus on what fits and discount what does not and if they fail to consider the possibility that equally accurate descriptions can be produced if their minds are not being read (Fischhoff & Beyth-Marom, 1983; Forer, 1949; Hyman, 1977). People who wish to believe in astrology or the predictive power of psychics will have no problem finding some predictions that have turned out to be true, and this may suffice to strengthen their belief if they fail to consider either predictions that proved not to be accurate or the possibility that people without the ability to see the future could make predictions with equally high (or low) hit rates. A confirmation

bias can work here in two ways: (a) people may attend selectively to what is said that turns out to be true, ignoring or discounting what turns out to be false, and (b) they may consider only $p(D|H)$, the probability that what was said would be said if the seer could really see, and fail to consider $p(D|\sim H)$, the probability that what was said would be said if the seer had no special psychic powers. The tendency of gamblers to explain away their losses, thus permitting themselves to believe that their chances of winning are higher than they really are, also illustrates the overweighting of supportive evidence and the underweighting of opposing evidence (Gilovich, 1983).

Seeing what one is looking for. People sometimes see in data the patterns for which they are looking, regardless of whether the patterns are really there. An early study by Kelley (1950) of the effect of expectations on social perception found that students' perceptions of social qualities (e.g., relative sociability, friendliness) of a guest lecturer were influenced by what they had been led to expect from a prior description of the individual. Forer (1949) demonstrated the ease with which people could be convinced that a personality sketch was an accurate depiction of themselves and their disinclination to consider how adequately the sketch might describe others as well.

Several studies by Snyder and his colleagues involving the judgment of personality traits lend credence to the idea that the degree to which what people see or remember corresponds to what they are looking for exceeds the correspondence as objectively assessed (Snyder, 1981, 1984; Snyder & Campbell, 1980; Snyder & Gangestad, 1981; Snyder & Swann, 1978a, 1978b). In a study representative of this body of work, participants would be asked to assess the personality of a person they are about to meet. Some would be given a sketch of an extrovert (sociable, talkative, outgoing) and asked to determine whether the person is that type. Others would be asked to determine whether the person is an introvert (shy, timid, quiet). Participants tend to ask questions that, if given positive answers, would be seen as strongly confirmatory and that, if given negative answers, would be weakly disconfirmatory of the personality type for which they are primed to look (Snyder, 1981; Snyder & Swann, 1978a). Some of the questions that people ask in these

situations are likely to evoke similar answers from extroverts and introverts (Swann, Giuliano, & Wegner, 1982)—answers not very diagnostic with respect to personality type. When this is the case, askers find it easy to see the answers they get as supportive of the hypothesis with which they are working, independently of what that hypothesis is.

The interpretation of the results of these studies is somewhat complicated by the finding of a tendency for people to respond to questions in a way that, in effect, acquiesces to whatever hypothesis the interrogator is entertaining (Lenski & Leggett, 1960; Ray, 1983; Schuman & Presser, 1981; Snyder, 1981; Zuckerman et al., 1995). For example, if the interrogator hypothesizes that the interviewee is an extrovert and asks questions that an extrovert would be expected to answer in the affirmative, the interviewee is more likely than not to answer in the affirmative independently of his or her personality type. Snyder, Tanke, and Berscheid (1977) have reported a related phenomenon. They found that male participants acted differently in a phone conversation toward female participants who had been described as attractive than toward those who had been described as unattractive, and that their behavior evoked more desirable responses from the "attractive" than from the "unattractive" partners.

Such results suggest that responders may inadvertently provide evidence for a working hypothesis by, in effect, accepting the assumption on which the questions or behavior are based and behaving in a way that is consistent with that assumption. Thus the observer's expectations become self-fulfilling prophecies in the sense suggested by Merton (1948, 1957); the expectations are "confirmed" because the behavior of the observed person has been shaped by them, to some degree. Studies in education have explored the effect of teachers' expectations on students' performance and have obtained similar results (Dusek, 1975; Meichenbaum, Bowers, & Ross, 1969; Rosenthal, 1974; Rosenthal & Jacobson, 1968; Wilkins, 1977; Zanna, Sheras, Cooper, & Shaw, 1975).

Darley and Fazio (1980) noted the importance of distinguishing between the case in which the behavior of a target person has changed in response to the perceiver's expectation-guided actions and that in which behavior has not changed but is interpreted to have done so in a

way that is consistent with the perceiver's expectations. Only the former is considered an example of a self-fulfilling prophecy. In that case, the change in behavior should be apparent to an observer whose observations are not subject to distortion by the expectations in question. When people's actions are interpreted in ways consistent with observers' expectations in the absence of any interaction between observer and observed, the self-fulfilling-prophecy effect cannot be a factor.

Darley and Gross (1983) provided evidence of seeing what one is looking for under the latter conditions noted above. These authors had two groups of people view the same videotape of a child taking an academic test. One of the groups had been led to believe that the child's socioeconomic background was high and the other had been led to believe that it was low. The former group rated the academic abilities, as indicated by what they could see of her performance on the test, as above grade level, whereas the latter group rated the same performance as below grade level. Darley and Gross saw this result as indicating that the participants in their study formed a hypothesis about the child's abilities on the basis of assumptions about the relationship between socioeconomic status and academic ability and then interpreted what they saw in the videotape so as to make it consistent with that hypothesis.

Numerous studies have reported evidence of participants seeing or remembering behavior that they expect. Sometimes the effects occur under conditions in which observers interact with the people observed and sometimes under conditions in which they do not. "Confirmed" expectations may be based on ethnic (Duncan, 1976), clinical (Langer & Abelson, 1974; Rosenhan, 1973; Swann et al., 1982), educational (Foster, Schmidt, & Sabatino, 1976; Rosenthal & Jacobson, 1968), socioeconomic (Rist, 1970), and lifestyle (Snyder & Uranowitz, 1978) stereotypes, among other factors. And they can involve induced expectations regarding oneself as well as others (Mischel, Ebbensen, & Zeiss, 1973).

Pennebaker and Skelton (1978) have pointed out how a confirmation bias can reinforce the worst fears of a hypochondriac. The body more or less continuously provides one with an assortment of signals that, if attended to, could be interpreted as symptoms of illness of one sort or another. Normally people ignore these signals. However, if one suspects that one is ill, one is likely to begin to attend to these signals and to notice those that are consistent with the assumed illness. Ironically, the acquisition of factual knowledge about diseases and their symptoms may exacerbate the problem. Upon learning of a specific illness and the symptoms that signal its existence, one may look for those symptoms in one's own body, thereby increasing the chances of detecting them even if they are not out of the normal range (Woods, Matterson, & Silverman, 1966).

Similar observations apply to paranoia and a variety of neurotic or psychotic states. If one believes strongly that one is a target of other people's ill will or aggression, one is likely to be able to fit many otherwise unaccounted-for incidents into this view. As Beck (1976) has suggested, the tendency of people suffering from mental depression to focus selectively on information that gives further reason to be depressed and ignore information of a more positive nature could help perpetuate the depressed state.

The results of experiments using such abstract tasks as estimating the proportions of beads of different colors in a bag after observing the color(s) of one or a few beads shows that data can be interpreted as favorable to a working hypothesis even when the data convey no diagnostic information. The drawing of beads of a given color may increase one's confidence in a hypothesis about the color distribution in the bag even when the probability of drawing a bead of that color is the same under the working hypothesis and its complement (Pitz, 1969; Troutman & Shanteau, 1977). After one has formed a preference for one brand of a commercial product over another, receiving information about an additional feature that is common to the two brands may strengthen one's preexisting preference (Chernev, 1997). Similarly, observers of a sports event may describe it very differently depending on which team they favor (Hastorf & Cantril, 1954).

It is true in science as it is elsewhere (Mitroff, 1974) that what one sees—actually or metaphorically—depends, to no small extent, on what one looks for and what one expects. Anatomist A. Kölliker criticized Charles Darwin for having

advanced the cause of teleological thinking, and botanist Asa Gray praised Darwin for the same reason. Meanwhile, biologist T. H. Huxley and naturalist Ernst Haeckel praised Darwin for thoroughly discrediting thinking of this kind (Gigerenzer et al., 1989).

Several investigators have stressed the importance of people's expectations as sources of bias in their judgments of covariation (Alloy & Abramson, 1980; Alloy & Tabachnik, 1984; Camerer, 1988; Crocker, 1981; Golding & Rorer, 1972; Hamilton, 1979; D. Kuhn, 1989; Nisbett & Ross, 1980). The belief that two variables are related appears to increase the chances that one will find evidence consistent with the relationship and decrease the chances of obtaining evidence that tends to disconfirm it. Judgments of covariation tend to be more accurate when people lack strong preconceptions of the relationship between the variables of interest or when the relationship is consistent with their preconceptions than when they have preconceptions that run counter to the relationship that exists.

The perception of a correlation where none exists is sometimes referred to as an *illusory correlation*. The term also has been applied when the variables in question are correlated but the correlation is perceived to be higher than it actually is Chapman and Chapman (1967a, 1967b, 1969; see also Chapman, 1967) did early work on the phenomenon. These investigators had participants view drawings of human figures, each of which occurred with two statements about the characteristics of the person who drew it. Statements were paired with drawings in such a way as to ensure no relationship between drawings and personality types, nevertheless participants saw the relationships they expected to see.

Nisbett and Ross (1980) summarized the Chapmans' work on illusory correlation by saying that "reported covariation was shown to reflect true covariation far *less* than it reflected theories or preconceptions of the nature of the associations that 'ought' to exist" (p. 97). Goldberg (1968) said that this research "illustrates the ease with which one can 'learn' relationships which do *not* exist" (p. 493). For present purposes, it provides another illustration of how the confirmation bias can work: the presumption of a relationship predisposes one to

find evidence of that relationship, even when there is none to be found or, if there is evidence to be found, to overweight it and arrive at a conclusion that goes beyond what the evidence justifies.

A form of stereotyping involves believing that specific behaviors are more common among people who are members of particular groups than among those who are not. There is a perceived correlation between group membership and behavior. Such perceived correlations can be real or illusory. One possible explanation of the perception of illusory correlations is that unusual behavior by people in distinctive groups is more salient and easily remembered than similar behavior by people who are not members of those groups (Feldman, Camburn, & Gatti, 1986; Hamilton, Dugan, & Trolier, 1985). Another possibility is that, once a person is convinced that members of a specific group behave in certain ways, he or she is more likely to seek and find evidence to support the belief than evidence to oppose it, somewhat independently of the facts.
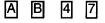
Some investigators have argued that people typically overestimate the degree to which behavior in different situations can be predicted from *trait variables* or the degree to which an individual's typical social behavior can be predicted from a knowledge of that person's behavior on a given occasion (Kunda & Nisbett, 1986). An *illusion of consistency*, according to this view, leads people to misjudge the extent to which a friendly individual's behavior is consistently friendly or a hostile individual's behavior is consistently hostile (Jennings, Amabile, & Ross, 1982; Mischel, 1968; Mischel & Peake, 1982). It is easy to see how a pervasive confirmation bias could be implicated in this illusion: once one has categorized an individual as friendly or hostile, one will be more attuned to evidence that supports that categorization than to evidence that undermines it.

A more subtle problem relating to categorization involves the phenomenon of *reification*. Taxonomies that are invented as conceptual conveniences often come to be seen as representing the way the world is really structured. Given the existence of a taxonomy, no matter how arbitrary, there is a tendency to view the world in terms of the categories it provides. One tends to

fit what one sees into the taxonomic bins at hand. In accordance with the confirmation bias, people are more likely to look for, and find, confirmation of the adequacy of a taxonomy than to seek and discover evidence of its limitations.

Formal Reasoning and the Selection Task

I have already considered a task invented by Wason and much used in rule-discovery experiments that revealed a tendency for people to test a hypothesized rule primarily by considering instances that are consistent with it. Wason (1966, 1968) invented another task that also has been widely used to study formal reasoning. In a well-known version of this task, participants see an array of cards and are told that each card has a letter on one side and a number on the other. Each of the cards they see shows either a vowel, a consonant, an even number, or an odd number, and participants are asked to indicate which cards one would have to turn over in order to determine the truth or falsity of the following statement: *If a card has a vowel on one side then it has an even number on the other side.*

Suppose the array that people performing this task see is as follows:

$$\boxed{A} \quad \boxed{B} \quad \boxed{4} \quad \boxed{7}$$

Given this set of cards, experimenters have generally considered selection of those showing A and 7 to be correct, because finding an odd number on the other side of the A or finding a vowel behind the 7 would reveal the statement to be false. The cards showing B and 4 have been considered incorrect selections, because whatever is on their other sides is consistent with the statement. In short, one can determine the claim to be false by finding *either* the card showing the A or the card showing the 7 to be inconsistent with it, or one can determine the claim to be true by finding *both* of these cards to be consistent with it. Wason found that people performing this task are most likely to select only the card showing a vowel or the card showing a vowel and the one showing an even number; people seldom select either the card showing a consonant or the one showing an odd number. Numerous investigators have obtained essentially the same result. Experiments with this task and variants of it have been

reviewed many times (Cosmides, 1989; Evans, 1982; Evans, Newstead, & Byrne, 1993; Tweney & Doherty, 1983; Wason & Johnson-Laird, 1972).

The logic of Wason's selection task is that of the conditional: *if P then Q*. In the case of the above example, P is *there is a vowel on one side*, and Q is *there is an even number on the other*. Selecting the card showing the 7 is analogous to checking to see if the *not-Q* case is accompanied by *not-P*, as it must be if the conditional is true. The basic finding of experiments with the task lends support to the hypothesis—which is supported also by other research (Evans, Newstead, & Byrne, 1993; Hollis, 1970)—that people find the *modus tollens* argument (*not-Q*, therefore *not-P*) to be less natural than the *modus ponens* form (*P*, therefore *Q*). And it is consistent with the idea that, given the objective of assessing the credibility of a conditional assertion, people are more likely to look for the presence of the consequent given the presence of the antecedent than for the absence of the antecedent given the absence of the consequent.

Several experiments have shown that performance of the selection task tends to be considerably better when the problem is couched in familiar situational terms rather than abstractly (Johnson-Laird, Legrenzi, & Legrenzi, 1972; Wason & Shapiro, 1971), although it is by no means always perfect in the former case (Einhorn & Hogarth, 1978). People also generally do better when the task is couched in terms that require *deontic reasoning* (deciding whether a rule of behavior—e.g., permission, obligation, promise—has been violated) rather than *indicative reasoning* (determining whether a hypothesis is true or false; Cheng & Holyoak, 1985; Cosmides, 1989; Cosmides & Tooby, 1992; Gigerenzer & Hug, 1992; Griggs & Cox, 1993; Kroger, Cheng, & Holyoak, 1993; Manktelow & Over, 1990, 1991; Markovits & Savary, 1992; Valentine, 1985; Yachanin & Tweney, 1982).

In relating confirmation bias to Wason's selection task, it is important to make three distinctions. The first is the distinction between the objective of specifying which of four cards in view must be turned over in order to determine the truth or falsity of the assertion *with respect to those four cards* and the objective of saying which of the four *types* of cards represented should be turned over in order to determine the plausibility of the assertion

more generally. I believe that in most of the earlier experiments, the first of these tasks was the intended one, although instructions have not always made it explicit that this was the case. The distinction is critical because what one should do when given the first objective is clear and not controversial. However the answer to what one should do when given the second objective is considerably more complicated and debatable.

When the task is understood in the first way, the only correct answer is the card showing the vowel and the one showing the odd number. However, when one believes one is being asked how one should go about obtaining evidence regarding the truth or falsity of an assertion of the form *if P then Q* where P and Q can be understood to be proxies for larger sets, it is no longer so clear that the cards showing P and $\sim Q$ are the best choices. Depending on the specifics of the properties represented by P and Q and what is known or assumed about their prevalence in the population of interest, P may be the only selection that is likely to be very informative, and selection of $\sim Q$ may be essentially pointless (Kirby, 1994b; Nickerson, 1996; Oaksford & Chater, 1994; Over & Evans, 1994).

Although experimenters have often, perhaps more often than not, intended that the selection task be interpreted in the first of the two ways described, the second interpretation seems to me more representative of the kind of conditional reasoning that people are required to do in everyday life; it is hard to think of everyday examples of needing to determine the truth or falsity of a claim about four entities like the cards in the selection task, all of which are immediately available for inspection. Perhaps a tendency to carry over to the task, uncritically, an effective approach in the world of everyday hypothesis testing that is not always effective in the contrived world of the laboratory contributes to performance on the selection task.

The second and third distinctions pertain to instances when the task is understood to involve only the four cards in view. They are the distinctions that were made in the context of the discussion of how confirmation bias relates to "find-the-rule" tasks and other tasks involving conditional syllogisms, namely the distinction between evidence that is confirmatory by virtue of the presence of two entities rather than by their joint absence and the distinction between

logical and psychological confirmation. As applied to the selection task, confirmation bias has typically connoted a tendency to look for the joint occurrence of the items specified in the task description. In the original version of the task, this means looking for instances of the joint occurrence of a vowel and an even number, and thus the turning over of the cards showing A and 4, inasmuch as these are the only cards of the four shown that could have both of the named properties.

Selection of evidence that can only be confirmatory of the hypothesis under consideration—or avoidance of evidence that could possibly falsify the hypothesis—is one interpretation of the confirmation bias that is not consistent with the selection of A and 4 in Wason's task. A is a potentially falsifying card; if the rule that cards that have a vowel on one side have an even number on the other is false, turning over the card showing A could reveal the fact. So in selecting A one is not avoiding the possibility of falsification; one is performing an appropriate test, even in the strictest Popperian sense. The only way to avoid the possibility of disconfirming the hypothesized rule in Wason's task is to select either or both of the cards showing B and 4, inasmuch as nothing that is on the other sides of these cards could be inconsistent with it.

The fact that, when given the selection task, people are at least as likely to select A as they are to select 4 weighs against the idea that their behavior can be attributed to a logically based desire to seek only confirming evidence and to avoid potentially disconfirming evidence. I say *logically based* because it may be that people select A because they are seeking confirming evidence, even though their selection also provides the opportunity to acquire falsifying evidence. From a strictly logical point of view, although selecting A is part of the correct solution, selecting only A, or A and 4, not only fails to ensure the discovery that the hypothesized rule is false if it is false, it also fails to establish its truth if it is true. The only way to demonstrate its truth is to rule out the possibility that it is false, and this means checking all (both) cases (A and 7) that could possibly show it to be false if it is.

So again, behavior that has been interpreted as evidence of a confirmation bias is not strongly confirmatory in a logical sense, al-

though it may well be seen (erroneously) as so by those who display it. Inasmuch as an observation can be confirming in a psychological sense (i.e., interpreted as confirming evidence) independently of whether it is logically confirmatory, perhaps the confirmation bias should be thought of as a tendency to seek evidence that increases one's confidence in a hypothesis regardless of whether it should. The distinction between logical confirmation and psychological confirmation deserves greater emphasis than it has received.

Although researchers have interpreted the considerable collection of results that have been obtained from experiments with variations of the selection task as generally reflective of a confirmation bias, alternative hypotheses have also been proposed. L. J. Cohen (1981) argued that it is not necessary to suppose that people do anything more in testing a conditional rule than check whether the consequent holds true whenever the antecedent does. The fact that people sometimes select the card representing Q (4 in the case of the present example) as well as the one representing P (A) is because they falsely perceive *if P then Q* to be equivalent to *if Q then P* (*conversion error*); and the fact that they do not select the card representing *not-Q* (7 in the present example) is attributed to a failure to apply the law of the contrapositive.

Several investigators have argued that people interpret the conditional relationship *if P then Q* as equivalent to the biconditional relationship *if and only if P then Q,* or that they fail to distinguish between *if P then Q* and *if Q then P,* especially as the *if–then* construction is used in everyday language (Chapman & Chapman, 1959; L. J. Cohen, 1981; Henle, 1962; Legrenzi, 1970; Politzer, 1986; Revlis, 1975a, 1975b). Others have assumed the operation of a *matching bias* whereby people tend to focus on (and select) cards explicitly named in the statement of the rule to be tested (Evans, 1972; Evans & Lynch, 1973; Hoch & Tschirgi, 1983). Evans (1989) has suggested that people make their selections, without thinking much about them, on the basis of a "preconscious heuristic judgment of relevance, probably linguistically determined" (p. 108).

Recent theorizing about the selection task has given rise to several new explanations of how people interpret it and deal with it, especially as presented in various concrete frames of refer-

ence. An idea common to several of these explanations is that people's performance of the task is indicative of behavior that has proved to be adaptively effective in analogous real-world situations. Cosmides (1989), for example, has argued that the typical results are consistent with the idea that people are predisposed to look for cheating on a social contract and are particularly skilled at detecting it. Consequently, they do well on selection tasks in which checking the *not-Q* condition is analogous to looking for an instance of violating a tacit social agreement, like failing to meet an obligation. Gigerenzer and Hug (1992) have defended a similar point of view.

Liberman and Klar (1996) proposed an alternative explanation of the results of research on cheating detection. They argued that a cheating-detection perspective is neither necessary nor sufficient to yield the P and *not-Q* selection; they contend that the logically correct choice will be likely if the conditional *if P then Q* is understood to represent a deterministic (as opposed to a probabilistic) unidirectional relationship between P and Q, if what constitutes a violation of the rule is clear, and if it is understood that the task is to look for such a violation.

Other investigators have begun to view the selection task as more a problem of data selection and decision making than of logical inference (Kirby, 1994a, 1994b; Manktelow & Over, 1991, 1992). Oaksford and Chater (1994), for example, have proposed a model of the task according to which people should make selections so as to maximize their gain in information regarding the tenability of the hypothesis that the conditional *if P then Q* is true relative to the tenability that it is false. I have done an analysis that leads to a similar conclusion, at least when the task is perceived as that of deciding whether the conditional is true in a general sense, as opposed to being true of four specific cards (Nickerson, 1996).

Wason's selection task has proved to be one of the most fertile paradigms in experimental psychology. It would be surprising if all the results obtained with the many variations of the task could be accounted for by a single simple hypothesis. I believe that, taken together, the results of experimentation support the hypothesis that one of the several factors determining performance on this task is a confirmation bias

that operates quite pervasively. This is not to suggest that a confirmation bias explains all the findings with this task, or even that it is the most important factor in all cases, but only that it exists and often plays a substantive role.

The Primacy Effect and Belief Persistence

When a person must draw a conclusion on the basis of information acquired and integrated over time, the information acquired early in the process is likely to carry more weight than that acquired later (Lingle & Ostrom, 1981; Sherman, Zehner, Johnson, & Hirt, 1983). This is called the *primacy effect*. People often form an opinion early in the process and then evaluate subsequently acquired information in a way that is partial to that opinion (N. H. Anderson & Jacobson, 1965; Jones & Goethals, 1972; Nisbett & Ross, 1980; Webster, 1964). Francis Bacon (1620/1939) observed this tendency centuries ago: "the first conclusion colors and brings into conformity with itself all that come after" (p. 36).

Peterson and DuCharme (1967) had people sample a sequence of colored chips and estimate the probability that the sequence came from an urn with a specified distribution of colors rather than from a second urn with a different distribution. The sampling was arranged so that the first 30 trials favored one urn, the second 30 favored the other, and after 60 trials the evidence was equally strong for each possibility. Participants tended to favor the urn indicated by the first 30 draws, which is to say that the evidence in the first 30 draws was not countermanded by the evidence in the second 30 even though statistically it should have been.

Bruner and Potter (1964) showed people the same picture on a series of slides. The first slide was defocused so the picture was not recognizable. On successive slides the focus was made increasingly clear. After each presentation the participant was asked to identify what was shown. A hypothesis formed on the basis of a defocused picture persisted even after the picture was in sufficiently good focus that participants who had not looked at the poorly focused picture were able to identify it correctly. Jones, Rock, Shaver, Goethals, and Ward (1968) had participants form opinions about people's problem-solving abilities by watching them in action. The problem solvers were judged to be more competent if they solved many problems early in a problem series and few late than if they did the reverse.

The primacy effect is closely related to (and can perhaps be seen as a manifestation of) *belief persistence*. Once a belief or opinion has been formed, it can be very resistive to change, even in the face of fairly compelling evidence that it is wrong (Freedman, 1964; Hayden & Mischel, 1976; Luchins, 1942, 1957; Rhine & Severance, 1970; Ross, 1977; Ross & Lepper, 1980; Ross, Lepper, & Hubbard, 1975). Moreover it can bias the evaluation and interpretation of evidence that is subsequently acquired. People are more likely to question information that conflicts with preexisting beliefs than information that is consistent with them and are more likely to see ambiguous information to be confirming of preexisting beliefs than disconfirming of them (Ross & Anderson, 1982). And they can be quite facile at explaining away events that are inconsistent with their established beliefs (Henrion & Fischhoff, 1986).

I have already discussed studies by Pitz (1969) and Troutman and Shanteau (1977) showing that people sometimes take normatively uninformative data as supportive of a favored hypothesis. One could interpret results from these studies as evidence of a belief-preserving bias working in situations in which the "information" in hand is not really informative. This is in keeping with the finding that two people with initially conflicting views can examine the same evidence and both find reasons for increasing the strength of their existing opinions (Lord, Ross, & Lepper, 1979). The demonstration by Pitz, Downing, and Reinhold (1967) that people sometimes interpret evidence that should count against a hypothesis as counting in favor of it may be seen as an extreme example of the confirmation bias serving the interest of belief preservation.

Ross and his colleagues have also shown experimentally that people find it extremely easy to form beliefs about or generate explanations of individuals' behavior and to persevere in those beliefs or explanations even after learning that the data on which the beliefs or explanations were originally based were fictitious (Ross et al., 1975; Ross, Lepper, Strack, & Steinmetz, 1977). Ross et al. (1975), for example, had people attempt to distinguish between authentic and unauthentic suicide

notes. As the participants made their choices, they were given feedback according to a predetermined schedule that was independent of the choices they made, but that insured that the feedback some participants received indicated that they performed far above average on the task while that which others received indicated that they performed far below average.

Following completion of Ross et al.'s (1975) task, researchers informed participants of the arbitrary nature of the feedback and told them that their rate of "success" or "failure" was predetermined and independent of their choices. When the participants were later asked to rate their ability to make such judgments, those who had received much positive feedback on the experimental task rated themselves higher than did those who had received negative feedback, despite being told that they had been given arbitrary information. A follow-up experiment found similar perseverance for people who observed others performing this task (but did not perform it themselves) and also observed the debriefing session.

Nisbett and Ross (1980) pointed out how a confirmation bias could contribute to the perseverance of unfounded beliefs of the kind involved in experiments like this. Receiving feedback that supports the assumption that one is particularly good or particularly poor at a task may prompt one to search for additional information to confirm that assumption. To the extent that such a search is successful, the belief that persists may rest not exclusively on the fraudulent feedback but also on other evidence that one has been able to find (selectively) in support of it.

It is natural to associate the confirmation bias with the perseverance of false beliefs, but in fact the operation of the bias may be independent of the truth or falsity of the belief involved. Not only can it contribute to the perseverance of unfounded beliefs, but it can help make beliefs for which there is legitimate evidence stronger than the evidence warrants. Probably few beliefs of the type that matter to people are totally unfounded in the sense that there is no legitimate evidence that can be marshalled for them. On the other hand, the data regarding confirmation bias, in the aggregate, suggest that many beliefs may be held with a strength or degree of certainty that exceeds what the evidence justifies.

Own-Judgment Evaluation

Many researchers have done experiments in which people have been asked to express their degree of confidence in judgments that they have made. When participants have expressed confidence as probability estimates or as ratings that, with some plausible assumptions, can be transformed into probability estimates, it has been possible to compare confidence with performance on the primary task. Thus researchers can determine for each confidence judgment the percentage of the correct items on the primary task to which that judgment was assigned. Plots of *actual percent correct* against *percent correct "predicted" by the confidence judgments* are often referred to as *calibration curves;* perfect calibration is represented by the unit line, which indicates that for a given confidence level, X, the proportion of all the judgments with that level that were correct was X.

In general, people tend to express a higher degree of confidence than is justified by the accuracy of their performance on the primary task, which is to say that calibration studies have typically shown overconfidence to be more common than underconfidence (Einhorn & Hogarth, 1978; Fischhoff, 1982; Lichtenstein & Fischhoff, 1977; Lichtenstein, Fischhoff, & Phillips, 1977; Pitz, 1974; Slovic, Fischhoff, & Lichtenstein, 1977). Kahneman and Tversky (1973) refer to the confidence that people feel for highly fallible performance as the *illusion of validity.* Being forced to evaluate one's views, especially when that includes providing reasons against one's position, has reduced overconfidence in some instances (Fischhoff, 1977; Hoch, 1984, 1985; Koriat, Lichtenstein, & Fischhoff, 1980; Tetlock & Kim, 1987). But generally overconfidence has only been reduced, not eliminated, and providing reasons against one's position is not something that most people do spontaneously.

One explanation of overconfidence starts with the assumption that people tend to be good judges of their knowledge as it relates to situations they are likely to encounter in everyday life. This explanation of overconfidence also notes that a minimum requirement for observing good calibration in experimental situations is that the questions people are to answer and that one to be used to judge the probability of the correctness of their answers

must be selected in such a way as to ensure that valid cues in the natural environment remain valid in the experimental situation. Juslin (1993) has argued that certain strategies commonly used to select items in experiments more or less ensure that valid knowledge from the participants' natural environment will be less valid in the experimental situation. More specifically, the argument is that such selection strategies typically result in sets of items for which cues leading to wrong answers are overrepresented relative to their commonness in the natural environment. Experiments showing good calibration for items selected at random from a set assumed to be representative of the natural environment (Gigerenzer, Hoffrage, & Kleinbölting, 1991; Juslin, 1993, 1994) support this explanation.

I believe this argument to be a forceful one. The evidence is fairly strong that some of the overconfidence reported, especially in experiments that require people to answer general-knowledge questions in a forced-choice format, is an artifact of item-selection procedures that do not ensure that cues have the same validity in experimental situations that they have in typical real-life situations. Not all studies of confidence or calibration have used general-knowledge questions and the forced-choice format for answers, however, and overconfidence has also been observed in other contexts. Researchers have shown that assessments of peoples' ability to recognize faces, for example, correlate poorly with performance on a face-recognition task (Baddeley & Woodhead, 1983) and observed that retrospective judgments of comprehension of expository texts are higher than justified (Glenberg, Wilkinson, & Epstein, 1982).

Experts are not immune from the illusion of validity. Attorneys tend, in the aggregate, to express confidence of obtaining outcomes better than those they actually obtain (Loftus & Wagenaar, 1988; Wagenaar & Keren, 1986). Other professionals who have been found to be overconfident when making judgments in their own areas of expertise include physicians (Lusted, 1977), psychologists (Oskamp, 1965), and engineers (Kidd, 1970). Experts appear to do better when there is a reliable basis for statistical prediction, such as when predicting bridge hands (Keren, 1987) or betting odds for horse racing (Hausch, Ziemba, & Rubenstein, 1981). On the other hand, despite the fact that clinical diagnoses based on case statistics tend to be more accurate than those based on clinical judgments, clinicians typically have greater confidence in their own judgments than in those derived statistically from incidence data (Arkes, Dawes, & Christensen, 1986; Goldberg, 1968; Meehl, 1954, 1960; Sawyer, 1966). In contrast, professional weather forecasters tend to be relatively well calibrated, at least with respect to their weather predictions (Murphy & Winkler, 1974, 1977; Winkler & Murphy, 1968); this has been attributed to the fact that they receive constant and immediate feedback regarding the accuracy of their predictions, which is the kind of information that makes learning feasible.

Griffin and Tversky (1992) have made a distinction between *strength* (extremeness) of evidence and *weight* (predictive validity) of evidence. They hypothesized that people tend to focus primarily on strength and make some (typically insufficient) adjustments in response to weight. This hypothesis leads to the expectation that overconfidence will be the rule especially when strength is high and weight low, whereas underconfidence becomes more likely when the opposite is the case. This account is consistent with the finding that overconfidence is the general rule because of the assumed greater importance attached to evidentiary strength. The account is even more predictive of general overconfidence if one assumes that strong evidence is more common than weighty evidence, as Griffin and Tversky use these terms.

Another account of why people tend to be overconfident of their own knowledge is that when one has produced a tentative answer to a question, one's natural inclination is to search memory for evidence to support that answer and to fail to consider possible alternative answers (Graesser & Hemphill, 1991; Griffin, Dunning, & Ross, 1990; Hoch, 1985; Koriat et al., 1980; Shaklee & Fischhoff, 1982). This is similar to the explanation already mentioned of why people sometimes persevere with a belief even after learning that the information on which the belief was based was fictitious: after having formed the belief they sought and found independent data to corroborate it.

The Confirmation Bias in Real-World Contexts

As the foregoing review shows, the confirmation bias has been observed in a variety of guises

in many experimental situations. What makes it especially important to understand is that it can have significant consequences in many nonlaboratory contexts. The point is illustrated in what follows by a few examples.

Number Mysticism

Pythagoras discovered, by experimentation, how the pitch of a sound emitted by a vibrating string depends on the length of the string and was able to state this dependency in terms of simple numerical ratios: (a) two strings of the same material under the same tension differ in pitch by an octave when one of the strings is twice as long as the other, and (b) two strings the lengths of which are in the ratio 2 to 3 will produce a note and its fifth, and so on.

Observation was not the new thing that Pythagoras did in his study of harmonic relationships; people had been observing the heavens and recording what they saw for a long time. What he did that was new was manipulate what he was observing and take notice of the effects of those manipulations. He has been called the first experimentalist. Ironically, instead of establishing experimentation as an especially fruitful way to investigate the properties of the physical world, Pythagoras's discovery helped to usher in what Bell called "the golden age of number mysticism" and to delay the acceptance of experimentation as the primary method of science for 2000 years (Bell, 1946/1991). Pythagoras's intellectual heirs were so convinced of the validity of his pronouncement, "everything is number," that many of the most able thinkers over the next 2 millenia devoted much of their cognitive energy to the pursuit of numerology and the confirmation of its basic assumptions. It took a Galileo to kindle an interest in experimentation that would not again sputter and die.

Work done as recently as the mid-19th century involving the great pyramid of Egypt, which has extraordinary fascination for modern observers of artifacts of ancient cultures, illustrates the relevance of the confirmation bias to Pythagorean numerological pursuits. Much of this fascination is due to certain mathematical relationships discussed first by John Taylor (1859) and shortly later by Charles Smyth (1864/1890). Taylor noted, among other facts, that the ratio of twice the pyramid's base to its height was roughly the same as the ratio of the diameter of a circle to its circumference, which is to say, π. Smyth was inspired by Taylor's observations and set himself the task of discovering other mathematical relationships of interest that the monument might hide.

Smyth discovered that the ratio of the pyramid's base to the width of a casing stone was 365, the number of days in a year, and that the pyramid's height multiplied by 10^9 was approximately equal to the distance from the earth to the sun. By comparing pyramid length measurements in various ways, he was able to find numbers that correspond to many quantitative properties of the world that were presumably unknown when the pyramid was built. These include the earth's mean density, the period of precession of the earth's axis, and the mean temperature of the earth's surface. Von Däniken (1969) used the existence of such relationships as the basis for arguing that the earth had been visited by intelligent extraterrestrials in the past.

Gardner (1957) referred to Smyth's book as a classic of its kind illustrating beautifully "the ease with which an intelligent man, passionately convinced of a theory, can manipulate his subject matter in such a way as to make it conform to precisely held opinions" (p. 176). He pointed out that a complicated structure like the pyramid provides one with a great assortment of possible length measurements, and that anyone with the patience to juggle them is quite sure to find a variety of numbers that will coincide with some dates or figures that are of interest for historical or scientific reasons. One simply makes an enormous number of observations, tries every manipulation on measurements and measurement relationships that one can imagine, and then selects from the results those few that coincide with numbers of interest in other contexts. He wrote, "Since you are bound by no rules, it would be odd indeed if this search for Pyramid 'Truths' failed to meet with considerable success" (p. 177). The search for pyramid truths is a striking illustration of how a bias to confirm is expressed by selectivity in the search for and interpretation of information.

Witch Hunting

From a modern vantage point, the convictions and executions of tens of thousands of individu-

als for practicing witchcraft during the 15th, 16th, and 17th centuries in Western Europe and to a lesser extent in 18th-century New England, is a particularly horrific case of the confirmation bias functioning in an extreme way at the societal level. From the perspective of many, perhaps most, of the people of the time, belief in witchcraft was perfectly natural and sorcery was widely viewed as the reason for ills and troubles that could not otherwise be explained. Executioners meted out punishment for practising witchcraft—generally the stake or the scaffold—with appalling regularity. Mackay (1852/1932) put the number executed in England alone, during the first 80 years of the 17th century, at 40,000.

People believed so strongly in witchcraft that some courts had special rules of evidence to apply only in cases involving it. Mackay (1852/1932) quoted Bodinus, a 17th-century French authority, as follows:

> The trial of this offense must not be conducted like other crimes. Whoever adheres to the ordinary course of justice perverts the spirit of the law, both divine and human. He who is accused of sorcery should never be acquitted, unless the malice of the prosecutor be clearer than the sun; for it is so difficult to bring full proof of this secret crime, that out of a million witches not one would be convicted if the usual course were followed! (p. 528)

Torture was an accepted and widely practiced means of exacting confessions from accused persons. Failure to denounce a witch was in some places a punishable offense.

It is hard for those who live in a society that recognizes the principle that a person accused of a crime is to be presumed innocent until proven guilty beyond a reasonable doubt to imagine what it must have been like to live at a time when—at least with respect to the accusation of witchcraft—just the opposite principle held. On the other hand, it is perhaps too easy to assume that nothing like the witchcraft mania could occur in our enlightened age. A moment's reflection on the instances of genocide or attempted genocide that have occurred in recent times should make people wary of any such assumption. These too can be seen as instances of what can happen when special rules of evidence are used to protect and support beliefs that individuals, groups or nations want, for whatever reasons, very much to hold. Awareness of a prevalent bias toward confirmation even under relatively ideal conditions of evidence evaluation should make people doubly cautious about the potential for disaster when circumstances encourage the bias to be carried to excess.

Policy Rationalization

Tuchman (1984) described a form of confirmation bias at work in the process of justifying policies to which a government has committed itself: "Once a policy has been adopted and implemented, all subsequent activity becomes an effort to justify it" (p. 245). In the context of a discussion of the policy that drew the United States into war in Vietnam and kept the U.S. military engaged for 16 years despite countless evidences that it was a lost cause from the beginning, Tuchman argued that once a policy has been adopted and implemented by a government, all subsequent activity of that government becomes focused on justification of that policy:

> Wooden headedness, the source of self deception is a factor that plays a remarkably large role in government. It consists in assessing a situation in terms of preconceived fixed notions while ignoring or rejecting any contrary signs. It is acting according to wish while not allowing oneself to be deflected by the facts. It is epitomized in a historian's statement about Philip II of Spain, the surpassing wooden head of all sovereigns: "no experience of the failure of his policy could shake his belief in its essential excellence." (p. 7)

Folly, she argued, is a form of self-deception characterized by "insistence on a rooted notion regardless of contrary evidence" (p. 209).

At the beginning of this article, I gave examples of bias in the use of information that would not be considered illustrative of the *confirmation bias,* as that term is used here. These examples involved intentional selectivity in the use of information for the conscious purpose of supporting a position. Another example is that of politicians taking note of facts that are consistent with positions they have taken while intentionally ignoring those that are inconsistent with them. This is not to suggest, however, that all selective use of information in the political arena is knowing and deliberate and that confirmation bias, in the sense of unwitting selectivity, is not seen here. To the contrary, I suspect that this type of bias is especially prevalent in situations that are inherently complex and ambiguous, which many political

situations are. In situations characterized by interactions among numerous variables and in which the cause–effect relationships are obscure, data tend to be open to many interpretations. When that is the case, the confirmation bias can have a great effect, and people should not be surprised to see knowledgeable, well-intentioned people draw support for diametrically opposed views from the same set of data.

Medicine

The importance of testing theories, hypotheses, speculations, or conjectures about the world and the way it works, by understanding their implications vis-a-vis observable phenomena and then making the observations necessary to check out those implications, was a common theme in the writings of the individuals who gave science its direction in the 16th, 17th, and 18th centuries. This spirit of empirical criticism had not dominated thinking in the preceding centuries but was a new attitude. Theretofore observation, and especially controlled experimentation, had played second fiddle to logic and tradition.

Consider, for example, the stagnated status of medical knowledge: "For 1500 years the main source of European physicians' knowledge about the human body was not the body itself . . . [but] the works of an ancient Greek physician [Galen]. 'Knowledge' was the barrier to knowledge. The classic source became a revered obstacle." Boorstin (1985, p. 344), who wrote these words, noted the irony of the fact that Galen himself was an experimentalist and urged others who wished to understand anatomy or medicine to become hands-on investigators and not to rely only on reading what others had said. But Galen's readers found it easier to rely on the knowledge he passed on to them than to follow his advice regarding how to extend it.

Although some physicians did "experiment" with various approaches to the treatment of different diseases and ailments, as Thomas (1979) pointed out, until quite recently, this experimentation left something to be desired as a scientific enterprise:

> Virtually anything that could be thought up for the treatment of disease was tried out at one time or another, and, once tried, lasted decades or even centuries before being given up. It was, in retrospect, the most frivolous and irresponsible kind of human

> experimentation, based on nothing but trial and error, and usually resulting in precisely that sequence. Bleeding, purging, cupping, the administration of infusions of every known plant, solutions of every known metal, every conceivable diet including total fasting, most of these based on the weirdest imaginings about the cause of disease, concocted out of nothing but thin air—this was the heritage of medicine up until a little over a century ago. It is astounding that the profession survived so long, and got away with so much with so little outcry. Almost everyone seems to have been taken in. (p. 159)

How is it that ineffective measures could be continued for decades or centuries without their ineffectiveness being discovered? Sometimes people got better when they were treated; sometimes they did not. And sometimes they got better when they were not treated at all. It appears that people's beliefs about the efficacy of specific treatments were influenced more strongly by those instances in which treatment was followed by recovery than in either those in which it was not or those in which it occurred spontaneously. A prevailing tendency to focus exclusively or primarily on positive cases— cases in which treatment was followed by recovery—would go a long way toward accounting for the fact that the discoveries that certain diseases have a natural history and people often recover from them with or without treatment was not made until the 19th century.

Fortunately such a tendency is no longer characteristic of medical science as a whole; but one would be hard pressed to argue that it no longer influences the beliefs of many people about the efficacy of various treatments of medical problems. Every practitioner of a form of pseudomedicine can point to a cadre of patients who will testify, in all sincerity, to having benefited from the treatment. More generally, people engage in specific behaviors (take a pill, rest, exercise, think positively) for the express purpose of bringing about a specific health-related result. If the desired result occurs, the natural tendency seems to be to attribute it to what was done for the purpose of causing it; considering seriously the possibility that the result might have been obtained in the absence of the associated "cause" appears not to come naturally to us but to have to be learned.

Medical diagnosis, as practiced today, has been the subject of some research. In looking for causes of illness, diagnosticians tend to generate one or a small set of hypotheses very early in the

diagnostic session (Elstein, Shulman, & Sprafka, 1978). The guidance that a hypothesis in hand represents for further information gathering can function as a constraint, decreasing the likelihood that one will consider an alternative hypothesis if the one in hand is not correct. Failure to generate a correct hypothesis has been a common cause of faulty diagnosis in some studies (Barrows, Feightner, Neufeld, & Norman, 1978; Barrows, Norman, Neufeld, & Feightner, 1977). A hypothesis in hand also can bias the interpretation of subsequently acquired data, either because one selectively looks for data that are supportive of the hypothesis and neglects to look for disconfirmatory data (Barrows et al., 1977) or because one interprets data to be confirmatory that really are not (Elstein et al., 1978).

Studies of physicians' estimates of the probability of specific diagnoses have often shown estimates to be too high (Christensen-Szalanski & Bushyhead, 1981/1988; DeSmet, Fryback, & Thornbury, 1979). Christensen-Szalanski and Bushyhead (1981/1988) had physicians judge, on the basis of a medical history and physical examination, the probability that patients at a walk-in clinic had pneumonia. Only about 20 percent of the patients who were judged to have pneumonia with a probability between .8 and .9 actually had pneumonia, as determined by chest X rays evaluated by radiologists. These investigators also got physicians to rate the various possible outcomes from such a diagnosis and found no difference between the values assigned to the two possible correct diagnoses and no difference between the values assigned to the possible incorrect ones. This suggests that the overly high estimates of probability of disease were not simply the result of a strong preference of false positive over false negative results.

Results from other studies suggest that physicians do not do very well at revising existing probability estimates to take into account the results of diagnostic tests (Berwick, Fineberg, & Weinstein, 1981; Cassells, Schoenberger, & Grayboys, 1978). A diagnostic finding that fails to confirm a favored hypothesis may be discounted on the grounds that, inasmuch as there is only a probabilistic relationship between symptoms and diseases, a perfect match is not to be expected (Elstein & Bordage, 1979).

Judicial Reasoning

In the context of judicial proceedings, an attempt is made to decouple the process of acquiring information from that of drawing conclusions from that information. Jurors are admonished to maintain open minds during the part of a trial when evidence is being presented (before the jury-deliberation phase); they are not supposed to form opinions regarding what the verdict should be until all the evidence has been presented and they have been instructed by the judge as to their decision task. During the jury-deliberation phase, the jury's task is to review and consider, collectively, the evidence that has been presented and to arrive, through discussion and debate, at a consensus on the verdict. They are to be careful to omit from consideration any evidence that came to light during the trial that was deemed inadmissible and ordered stricken from the record.

The admonition to maintain an open mind during the evidence-presentation phase of a trial seems designed to counter the tendency to form an opinion early in an evidence-evaluation process and then to evaluate further evidence with a bias that favors that initial opinion. Whether jurors are able to follow this admonition and to delay forming an opinion as to the truth or falsity of the allegations against the accused until the jury-deliberation phase is a matter of some doubt. It is at least a plausible possibility that individual jurors (and judges) develop their personal mental models of "what really happened" as a case develops and continuously refine and elaborate those models as evidence continues to be presented (Holstein, 1985). To the extent that this is the way jurors actually think, a model, as it exists at any point in time, may strongly influence how new information is interpreted. If one has come to believe that a defendant is guilty (or innocent), further evidence that is open to various interpretations may be seen as supportive of that belief. Opinions formed about a defendant on the basis of superficial cues (such as demeanor while giving testimony) can bias the interpretation of inherently ambiguous evidence (Hendry & Shaffer, 1989).

Results of mock-trial experiments indicate that, although there are considerable individual differences among mock jurors with respect to

how they approach their task (D. Kuhn et al., 1994), jurors often come to favor a particular verdict early in the trial process (Devine & Ostrom, 1985). The final verdicts that juries return are usually the same as the tentative ones they initially form (Kalven & Zeisel, 1966; Lawson, 1968). The results of some mock trials suggest that deliberation following the presentation of evidence tends to have the effect of making a jury's average predeliberation opinion more extreme in the same direction (Myers & Lamm, 1976).

The tendency to stick with initial tentative verdicts could exist because in most cases the initial conclusions stand up to further objective evaluation; there is also the possibility, however, that the tentative verdict influences jurors' subsequent thinking and biases them to look for, or give undo weight to, evidence that supports it. This possibility gains credence from the finding by Pennington and Hastie (1993) that participants in mock-jury trials were more likely to remember statements consistent with their chosen verdict as having been presented as trial evidence than statements that were inconsistent with this verdict. This was true both of statements that had been presented (hits) and of those that had not (false positives).

Science

Polya (1954a) has argued that a tendency to resist the confirmation bias is one of the ways in which scientific thinking differs from everyday thinking:

> The mental procedures of the trained naturalist are not essentially different from those of the common man, but they are more thorough. Both the common man and the scientist are led to conjectures by a few observations and they are both paying attention to later cases which could be in agreement or not with the conjecture. A case in agreement makes the conjecture more likely, a conflicting case disproves it, and here the difference begins: Ordinary people are usually more apt to look for the first kind of cases, but the scientist looks for the second kind. (p. 40)

If seeking data that would disconfirm a hypothesis that one holds is the general rule among scientists, the history of science gives us many exceptions to the rule (Mahoney, 1976, 1977; Mitroff, 1974). Michael Faraday was likely to seek confirming evidence for a hypothesis and ignore such disconfirming evidence as he obtained until the phenomenon

under study was reasonably well understood, at which time he would begin to pay more attention to disconfirming evidence and actively seek to account for it (Tweney & Doherty, 1983). Louis Pasteur refused to accept or publish results of his experiments that seemed to tell against his position that life did not generate spontaneously, being sufficiently convinced of his hypothesis to consider any experiment that produced counterindicative evidence to be necessarily flawed (Farley & Geison, 1974). When Robert Millikan published the experimental work on determining the electric charge of a single electron, for which he won the Nobel prize in physics, he reported only slightly more than half (58) of his (107) observations, omitting from publication those that did not fit his hypothesis (Henrion & Fischhoff, 1986).

It is not so much the critical attitude that individual scientists have taken with respect to their own ideas that has given science the success it has enjoyed as a method for making new discoveries, but more the fact that individual scientists have been highly motivated to demonstrate that hypotheses that are held by some other scientist(s) are false. The insistence of science, as an institution, on the objective testability of hypotheses by publicly scrutable methods has ensured its relative independence from the biases of its practitioners.

Conservatism among scientists. The fact that scientific discoveries have often met resistance from economic, technological, religious, and ideological elements outside science has been highly publicized. That such discoveries have sometimes met even greater resistance from scientists, and especially from those whose theoretical positions were challenged or invalidated by those discoveries, is no less a fact if less well known (Barber, 1961; Mahoney, 1976, 1977). Galileo, for example, would not accept Kepler's hypothesis that the moon is responsible for the tidal motions of the earth's oceans. Newton refused to believe that the earth could be much older than 6,000 years on the strength of the reasoning that led Archbishop Usher to place the date of creation at 4,004 BC. Huygens and Leibniz rejected Newton's concept of universal gravity because they could not accept the idea of a force extending throughout space not reducible to matter and motion.

Humphrey Davy dismissed John Dalton's ideas about the atomic structure of matter as

more ingenious than important. William Thomson (Lord) Kelvin, who died in 1907, some years after the revolutionary work of Joseph Thomson on the electron, never accepted the idea that the atom was decomposable into simpler components. Lev Landau was willing in 1932 to dismiss the laws of quantum mechanics on the grounds that they led to such a ridiculous prediction as the contraction of large burned out stars to essentially a point. Arthur Eddington rejected Subrahmanyan Chandrasekhar's prediction in the early 1930s that cold stars with a mass of more than about one half that of the sun would collapse to a point. Chandrasekhar was sufficiently discouraged by Eddington's reaction to leave off for the better part of his professional career the line of thinking that eventually led to the now widely accepted theory of black holes.

Theory persistence. The history of science contains many examples of individual scientists tenaciously holding on to favored theories long after the evidence against them had become sufficiently strong to persuade others without the same vested interests to discard them. I. B. Cohen (1985) has documented many of these struggles in his account of the role of revolution in science. All of them can be seen as examples of the confirmation bias manifesting itself as an unwillingness to give the deserved weight to evidence that tells against a favored view.

Roszak (1986) described the tenacity with which adherents to the cosmology of Ptolemy clung to their view in the face of mounting evidence of its untenability in this way:

> The Ptolemaic cosmology that prevailed in ancient times and during the Middle Ages had been compromised by countless contradictory observations over many generations. Still, it was an internally coherent, intellectually pleasing idea; therefore, keen minds stood by the familiar old system. Where there seemed to be any conflict, they simply adjusted and elaborated the idea, or restructured the observations in order to make them fit. If observations could not be made to fit, they might be allowed to stand along the cultural sidelines as curiosities, exceptions, freaks of nature. It was not until a highly imaginative constellation of ideas about celestial and terrestrial dynamics, replete with new concepts of gravitation, inertia, momentum, and matter, was created that the old system was retired. (p. 91)

Roszak pointed out also that scientists throughout the 18th and 19th centuries retained other inherited ideas in the fields of chemistry, geology, and biology in adjusted forms, despite increasing evidences of their inadequacies.

Science has held, often for very long times, some beliefs and theories that could have been invalidated if a serious effort had been made to show them to be false. The belief that heavier bodies fall faster than lighter ones, for example, prevailed from the time of Aristotle until that of Galileo. The experiments that Galileo performed to demonstrate that this belief was false could have been done at any time during that 2000-year period. This is a particularly interesting example of a persisting false belief because it might have been questioned strictly on the basis of reasoning, apart from any observations. Galileo posed a question that could have been asked by Aristotle or by anybody who believed that heavier bodies fall faster than lighter ones: If a 10 pound weight falls faster than a 1 pound weight, what will happen when the two are tied together? Will the 11 pound combination fall faster than the 10 pound weight, or will it fall more slowly because the 1 pound weight will hold back the 10 pound one?

Hawking (1988) argued that the fact that the universe is expanding could have been predicted from Newton's theory of gravity at any time after the late 17th century, and noted that Newton himself realized that the idea of universal gravity begged an explanation as to why the stars did not draw together. The assumption that the universe was static was a strong and persistent one, however. Newton dealt with the problem by assuming the universe was infinite and consequently had no center onto which it could collapse. Others introduced a notion that at very great distances gravity became repulsive instead of attractive. Einstein, in order to make his general theory of relativity compatible with the idea of a static universe, incorporated a "cosmological constant" in the theory which, in effect, nullified the otherwise expected contraction. Einstein later was embarrassed by this invention and saw it as his greatest mistake.

None of this is to suggest that scientists accept evidences of the inadequacy of an established theory with equanimity. Rather, I note that the typical reaction to the receipt of such evidences is not immediately to discard the theory to which the inadequacies relate but to find a way to defend it. The bias is definitely in the direction of giving the existing theory the

branded "eccentric" if not "obscurantist"; some of his most doubtful assumptions came to be regarded as so trivial that textbooks never even stated them. (p. 49, Footnote 1)

One can see a confirmation bias both in the difficulty with which new ideas break through opposing established points of view and in the uncritical allegiance they are often given once they have become part of the established view themselves.

Explanations of the Confirmation Bias

How is one to account for the confirmation bias and its prevalence in so many guises? Is it a matter of protecting one's ego, a simple reluctance to consider the possibility that a belief one holds or a hypothesis that one is entertaining is wrong? Is it a consequence of specific cognitive limitations? Does it reflect a lack of understanding of logic? Does it persist because it has some functional value? That is, does it provide some benefits that are as important as, or in some situations more important than, an attempt to determine the truth in an unbiased way would be?

The Desire to Believe

Philosophers and psychologists alike have observed that people find it easier to believe propositions they would like to be true than propositions they would prefer to be false. This tendency has been seen as one manifestation of what has been dubbed *the Pollyanna principle* (Matlin & Stang, 1978), according to which people are likely to give preferential treatment to pleasant thoughts and memories over unpleasant ones.

Finding a positive correlation between the probability that one will believe a proposition to be true and the probability that one will consider it to be desirable (Lefford, 1946; McGuire, 1960; Weinstein, 1980, 1989) does not, in itself, establish a causal link between desirability and perceived truth. The correlation could reflect a relationship between truth and desirability in the real world, whereby what is likely to be true is likely also to be desirable, and the same is conversely true. On the other hand, the evidence is strong that the correlation is the result, at least to some degree, of beliefs being influenced by preferences. The continuing susceptibility of

people to too-good-to-be-true promises of quick wealth is but one illustration of the fact that people sometimes demand very little in the way of compelling evidence to drive them to a conclusion that they would like to accept.

Although beliefs can be influenced by preferences, there is a limit to how much influence people's preferences can have. It is not the case, for most of us at least, that we are free to believe anything we want; what we believe must appear to us believable. We can be selective with respect to the evidence we seek, and we can tilt the scales when we weigh what we find, but we cannot completely ignore counterindicative evidence of which we are aware. Kunda (1990) has made this argument persuasively. The very fact that we sometimes seek to ignore or discount evidence that counts against what we would like to believe bears witness to the importance we attach to holding beliefs that are justified.

More generally, one could view, somewhat ironically perhaps, the tendency to treat data selectively and partially as a testament to the high value people attach to consistency. If consistency between beliefs and evidence were of no importance, people would have no reason to guard beliefs against data that are inconsistent with them. Consistency is usually taken to be an important requirement of rationality, possibly the most important such requirement. Paradoxically, it seems that the desire to be consistent can be so strong as to make it difficult for one to evaluate new evidence pertaining to a stated position in an objective way.

The quote from Mackay (1852/1932) that is used as an epigraph at the beginning of this article stresses the importance of motivation in efforts to confirm favored views. Some investigators have argued, however, that the basic problem is not motivational but reflects limitations of a cognitive nature. For example, Nisbett and Ross (1980) held that, on the whole, investigators have been too quick to attribute the behavior of participants in experimental situations to motivational biases when there were equally plausible alternative interpretations of the findings:

One wonders how strongly the theory of self-serving bias must have been held to prompt such uncritical acceptance of empirical evidence. . . . We doubt that careful investigation will reveal ego-enhancing or ego-defensive biases in attribution to be as pervasive or

potent as many lay people and most motivational
theorists presume them to be. (p. 233)

This argument is especially interesting in the
present context, because it invokes a form of
confirmation bias to account for the tendency of
some investigators to attribute certain behaviors
to motivational causes and to ignore what, in
Nisbett and Ross's view, are equally likely
alternative explanations.

The role of motivation in reasoning has been
a subject of debate for some time. Kunda (1990)
noted that many of the phenomena that once
were attributed to motivational variables have
been reinterpreted more recently in cognitive
terms; according to this interpretation, conclu-
sions that appear to be drawn only because
people want to draw them may be drawn
because they are more consistent with prior
beliefs and expectancies. She noted too that
some theorists have come to believe that
motivational effects are mediated by cognitive
processes. According to this view, "[p]eople
rely on cognitive processes and representations
to arrive at their desired conclusions, but
motivation plays a role in determining which of
these will be used on a given occasion" (Kunda,
1990, p. 480).

Kunda defended this view, arguing that the
evidence to date is consistent with the assump-
tion that motivation affects reasoning, but it
does so through cognitive strategies for access-
ing, constructing, and evaluating beliefs:

> Although cognitive processes cannot fully account for
> the existence of self-serving biases, it appears that they
> play a major role in producing these biases in that they
> provide the mechanisms through which motivation
> affects reasoning. Indeed, it is possible that motivation
> merely provides an initial trigger for the operation of
> cognitive processes that lead to the desired conclusions.
> (p. 493)

The primary cognitive operation hypoth-
esized to mediate motivational effects is the
biased searching of memory. Evidence of
various types converges, she argued, on the
conclusion that "goals enhance the accessibility
of those knowledge structures—memories, be-
liefs, and rules—that are consistent with desired
conclusions" (p. 494); "Motivation will cause
bias, but cognitive factors such as the available
beliefs and rules will determine the magnitude
of the bias" (p. 495).

Several of the accounts of confirmation bias
that follow stress the role of cognitive limita-

tions as causal factors. It is possible, however,
and probable, in my view, that both motivational
and cognitive factors are involved and that each
type can mediate effects of the other.

Information-Processing Bases for Confirmation Bias

The confirmation bias is sometimes attributed
in part to the tendency of people to gather
information about only one hypothesis at a time
and even with respect to that hypothesis to
consider only the possibility that the hypothesis
is true (or only the possibility that it is false) but
not to consider both possibilities simultaneously
(Tweney, 1984; Tweney & Doherty, 1983).
Doherty and Mynatt (1986) argued, for ex-
ample, that people are fundamentally limited to
think of only one thing at a time, and once
having focused on a particular hypothesis, they
continue to do so. This, they suggested, explains
why people often select nondiagnostic over
diagnostic information in Bayesian decision
situations. Suppose that one must attempt to
decide which of two diseases, A or B, a patient
with Symptoms X and Y has. One is informed of
the relative frequency of Symptom X among
people who have Disease A and is then given the
choice of obtaining either of the following items
of information: the relative frequency of people
with A who have Symptom Y or the relative
frequency of people with B who have symptom
X. Most people who have been given choices of
this sort opt for the first; they continue to focus
on the hypothesis that the patient has A, even
though learning the relative frequency of Y
given A does not inform the diagnosis, whereas
learning the relative frequency of X given B
does.

Assuming a restricted focus on a single
hypothesis, it is easy to see how that hypothesis
might become strengthened even if it is false. An
incorrect hypothesis can be sufficiently close to
being correct that it receives a considerable
amount of positive reinforcement, which may be
taken as further evidence of the correctness of
the hypothesis in hand and inhibit continued
search for an alternative. In many contexts
intermittent reinforcement suffices to sustain the
behavior that yields it.

People also can increase the likelihood of
getting information that is consistent with
existing beliefs and decrease the likelihood of

getting information that is inconsistent with them by being selective with respect to where they get information (Frey, 1986). The idea that people tend to expose themselves more to information sources that share their beliefs than to those that do not has had considerable credibility among social psychologists (Festinger, 1957; Klapper, 1960). Sears and Freedman (1967) have challenged the conclusiveness of much of the evidence that has been evoked in support of this idea. They noted that, when given a choice of information that is supportive of a view one holds and information that is supportive of an opposing view, people sometimes select the former and sometimes the latter, and sometimes they show no preference. Behavior in these situations seems to depend on a number of factors in addition to the polarity of the information with respect to one's existing views, such as people's level of education or social status and the perceived usefulness of the information that is offered.

Sears and Freedman (1967) stopped short, however, of concluding that people are totally unbiased in this respect. They noted the possibility that "dramatic selectivity in preferences may not appear at any given moment in time, but, over a long period, people may organize their surroundings in a way that ensures *de facto* selectivity" (p. 213). People tend to associate, on a long-term basis, with people who think more or less as they do on matters important to them; they read authors with whom they tend to agree, listen to news commentators who interpret current events in a way that they like, and so on. The extent to which people choose their associates because of their beliefs versus forming their beliefs because of their associates is an open question. But it seems safe to assume that it goes a bit in both directions. Finding lots of support for one's beliefs and opinions would be a natural consequence of principally associating with people with whom one has much in common.

Gilovich (1991) made the important point that for many beliefs or expectations confirmatory events are likely to be more salient than nonconfirmatory ones. If a fortune teller predicts several events in one's life sometime during the indefinite future, for example, the occurrence of any predicted event is more likely to remind one of the original prediction of that event than is its nonoccurrence. Events predicted by a fortune teller illustrate *one-sided events*. Unlike *two-sided events*, which have the characteristic that their nonoccurrence is as noticeable as their occurrence, one-sided events are likely to be noticed while their nonoccurrence is not. (An example of a two-sided event would be the toss of a coin following the prediction of a head. In this case the nonoccurrence of the predicted outcome would be as noticeable as its occurrence.)

Sometimes decision policies that rule out the occurrence of certain types of events preclude the acquisition of information that is counter-indicative with respect to a hypothesis. Consider, for example, the hypothesis that only students who meet certain admission requirements are likely to be successful as college students. If colleges admit only students who meet those requirements, a critical subset of the data that are necessary to falsify the hypothesis (the incidence of students who do not meet the admission requirements but are nevertheless successful college students) will not exist (Einhorn & Hogarth, 1978). One could argue that this preclusion may be justified if the hypothesis is correct, or nearly so, and if the negative consequences of one type of error (admitting many students who will fail) are much greater than those of the other type (failing to admit a few students who would succeed). But the argument is circular because it assumes the validity of the hypothesis in question.

Some beliefs are such that obtaining evidence that they are false is inherently impossible. If I believe, for example, that most crimes are discovered sooner or later, my belief may be reinforced every time a crime is reported by a law enforcement agency. But by definition, undiscovered crimes are not discovered, so there is no way of knowing how many or them there are. For the same reason, if I believe that most crimes go undiscovered, there is no way to demonstrate that this belief is wrong. No matter how many crimes are discovered, the number of undiscovered crimes is indeterminable because being undiscovered means being uncounted.

Some have also given information-processing accounts of why people, in effect, consider only the probability of an event, assuming the truth of a hypothesis of interest, and fail to consider the probability of the same event, assuming the falsity of that hypothesis. Evans (1989) argued

that one need not assume the operation of a *motivational bias*—a strong wish to confirm—in order to account for this failure. It could signify, according to Evans, a lack of understanding of the fact that, without a knowledge of $p(D|{\sim}H)$, $p(D|H)$ gives one no useful information about $p(H|D)$. That is, $p(D|H)$, by itself, is not diagnostic with respect to the truth or falsity of H. Possibly people confuse $p(D|H)$ with $p(H|D)$ and take its absolute value as an indication of the strength of the evidence in favor of H. Bayes's rule of inverse probability, a formula for getting from $p(D|H)$ to $p(H|D)$ presumably was motivated, in part, to resolve this confusion.

Another explanation of why people fail to consider alternatives to a hypothesis in hand is that they simply do not think to do so. Plausible alternatives do not come to mind. This is seen by some investigators to be, at least in part, a matter of inadequate effort, a failure to do a sufficiently extensive search for possibilities (Baron, 1985, 1994; Kanouse, 1972).

Positive-Test Strategy or Positivity Bias

Arguing that failure to distinguish among different senses of confirmation in the literature has contributed to misinterpretations of both empirical findings and theoretical prescriptions, Klayman and Ha (1987) have suggested that many phenomena of human hypothesis testing can be accounted for by the assumption of a general *positive-test strategy.* Application of this strategy involves testing a hypothesis either by considering conditions under which the hypothesized event is expected to occur (to see if it does occur) or by examining known instances of its occurrence (to see if the hypothesized conditions prevailed). The phenomenon bears some resemblance to the finding that, in the absence of compelling evidence one way or the other, people are more inclined to assume that a statement is true than to assume that it is false (Clark & Chase, 1972; Gilbert, 1991; Trabasso, Rollins, & Shaughnessy, 1971; Wallsten & González-Vallejo, 1994).

Baron, Beattie, and Hershey (1988), who referred to the positive-test strategy as the *congruence heuristic,* have shown that the tendency to use it can be reduced if people are asked to consider alternatives, but that they tend not to consider them spontaneously. The fact that people show a preference for questions that would yield a positive answer if the hypothesis is correct over questions that would yield a negative answer if the hypothesis is correct demonstrates that the positive-test strategy is not a simple consequence of the confirmation bias (Baron et al., 1988; Devine, Hirt, & Gehrke, 1990; Skov & Sherman, 1986).

One could view the results that Wason (1960) originally obtained with the number-triplet task, which constituted the point of departure for much of the subsequent work on confirmation bias, as a manifestation of the positive-test strategy, according to which one tests cases one thinks likely to have the hypothesized property (Klayman & Ha, 1987). As already noted, this strategy precludes discovering the incorrectness of a hypothesized rule when instances that satisfy the hypothesized rule constitute a subset of those that satisfy the correct one, but it is an effective strategy when the instances that satisfy the correct rule constitute a subset of those that satisfy the hypothesized one.

Suppose, for example, that the hypothesized rule is *successive even numbers* and the correct one is *increasing numbers.* All triplets that satisfy the hypothesized rule also satisfy the correct one, so using only such triplets as test cases will not reveal the incorrectness of the hypothesized rule. But if the hypothesized rule is *increasing numbers* and the correct rule is *successive even numbers,* the positive-test strategy—which, in this case means trying various triplets of increasing numbers—is likely to provide the feedback necessary to discover that the hypothesized rule is wrong. In both of the examples considered, the strategy is to select instances for testing that satisfy the hypothesized rule; it is effective in revealing the rule to be wrong in the second case, not because the tester intentionally selects instances that will show the rule to be wrong if it is wrong, but because the relationship between hypothesized and correct rules is such that the tester is likely to discover the hypothesized rule to be wrong by selecting test cases intended to show that it is right.

Klayman and Ha (1987) analyzed various possible relationships between the sets of triplets delimited by hypothesized and correct rules in addition to the two cases in which one set is a proper subset of the other—overlapping sets, disjoint sets, identical sets—and showed

that the likely effectiveness of positive-test strategy depends on which relationship pertains. They also analyzed corresponding cases in which set membership is probabilistic and drew a similar conclusion. With respect to disconfirmation, Klayman and Ha argued the importance of distinguishing between two strategies, one involving examination of instances that are expected not to have the target property, and the other involving examination of instances that one expects to falsify, rather than verify, the hypothesis. Klayman and Ha contended that failure to make this distinction clearly in the past has been responsible for some confusion and debate.

Several investigators have argued that a positive-test strategy should not necessarily be considered a biased information-gathering technique because questions prompted by this strategy generally do not preclude negative answers and, therefore, falsification of the hypothesis being tested (Bassock & Trope, 1984; Hodgins & Zuckerman, 1993; Skov & Sherman, 1986; Trope & Mackie, 1987). It is important to distinguish, however, between obtaining information from a test because the test was intended to yield the information obtained and obtaining information adventitiously from a test that was intended to yield something other than what it did. Consider again the number-triplet task. One might select, say, 6–7–8 for either of the following reasons, among others: (a) because one believes the rule to be *increasing numbers* and wants a triplet that fits the rule, or (b) because one believes the rule to be *successive even numbers* and wants to increase one's confidence in this belief by selecting a triplet that fails to satisfy the rule. The chooser expects the experimenter's response to the selected triplet to be positive in the first case and negative in the second. In either case, the experimenter's response could be opposite from what the chooser expects and show the hypothesized rule to be wrong; the response is adventitious because the test yielded information that the chooser was not seeking and did not expect to obtain.

The fact that people appear to be more likely than not to test a hypothesized rule by selecting cases that they believe will pass muster (that fit the hypothesized rule) and lend credence to the hypothesis by doing so justifies describing typical performance on find-the-rule tasks like the triplet task as illustrative of a confirmation bias. People appear to be much less likely to attempt to get evidence for a hypothesized rule by choosing a case that they believe does not fit it (and expecting an informative-negative response) or to select cases with the intended purpose of ruling out one or more currently plausible hypotheses from further consideration. When a test that was expected to yield an outcome that is positive with respect to a hypothesized rule does not do so, it seems appropriate to say that the positive-test strategy has proved to be adventitiously informative. Clearly, it is possible to select an item that is consistent with a hypothesized rule for the purpose of revealing an alternative rule to be wrong, but there is little evidence that this is often done. It seems that people generally select test items that are consistent with the rule they believe to be correct and seldom select items with falsification in mind.

Klayman and Ha (1987) argued that the positive-test strategy is sometimes appropriate and sometimes not, depending on situational variables such as the base rates of the phenomena of interest, and that it is effective under commonly occurring conditions. They argued too, however, that people tend to rely on it overly much, treating it as a default strategy to be used when testing must be done under less-than-ideal conditions, and that it accounts for many of the phenomena that are generally interpreted as evidence of a pervasive confirmation bias.

Evans (1989) has proposed an explanation of the confirmation bias that is similar in some respects to Klayman and Ha's (1987) account and also discounts the possibility that people intentionally seek to confirm rather than falsify their hypotheses. Cognitive failure, and not motivation, he argued, is the basis of the phenomenon:

> Subjects confirm, not because they want to, but because they cannot think of the way to falsify. The cognitive failure is caused by a form of selective processing which is very fundamental indeed in cognition—a bias to think about positive rather than negative information. (p. 42)

With respect to Wason's (1960) results with the number-triplet task, Evans suggested that rather than attempting to confirm the hypotheses they were entertaining, participants may simply have been unable to think of testing them in a

negative manner; from a logical point of view, they should have used potentially disconfirming test cases, but they failed to think to do so.

Evans (1989) cited the results of several efforts by experimenters to modify hypothesis-testing behavior on derivatives of Wason's number-set task by instructing participants about the importance of seeking negative or disconfirming information. He noted that although behavioral changes were sometimes induced, more often performance was not improved. This too was taken as evidence that results that have been attributed to confirmation bias do not stem from a desire to confirm, but rather from the difficulty people have in thinking in explicitly disconfirmatory terms.

Evans used the same argument to account for the results typically obtained with Wason's (1966, 1968) selection task that have generally been interpreted as evidence of the operation of a confirmation bias. Participants select named items in this task, according to this view, because these are the only ones that come to mind when they are thinking about what to do. The bias that is operating is not that of wanting to confirm a tentative hypothesis but that of being strongly inclined to think only of information that is explicitly provided in the problem statement. In short, Evans (1989) distinguished between confirmatory *behavior,* which he acknowledges, and confirmatory *intentions,* which he denies. The "demonstrable deficiencies in the way in which people go about testing and eliminating hypotheses," he contended, "are a function of selective processing induced by a widespread cognitive difficulty in thinking about any information which is essentially negative in its conception" (p. 63). The tendency to focus on positive information and fail to consider negative information is regarded not as a conscious cognitive strategy but as the result of preattentive processes. *Positive* is not synonymous with *confirmatory,* but as most studies have been designed, looking for positive cases is tantamount to looking for confirmation.

The idea that people have a tendency to focus more on positive than on negative information, like the idea of a confirmation bias, is an old one. An observation by Francis Bacon (1620/1939) can again illustrate the point: "It is the peculiar and perpetual error of the human understanding to be more moved and excited by affirmatives than negatives; whereas it ought

properly to hold itself indifferently disposed towards both alike" (p. 36). Evans's (1989) argument that such a bias can account, at least in part, for some of the phenomena that are attributed to a confirmation bias is an intuitively plausible one. A study in which Perkins et al., (1983) classified errors of reasoning made in informal arguments constructed by over 300 people of varying age and educational level supports the argument. Many of the errors Perkins et al. identified involved participants' failure to consider lines of reasoning that could be used to defeat or challenge their conclusions.

I will be surprised if a positivity bias turns out to be adequate to account for all of the ways in which what has here been called a *confirmation bias* manifests itself, but there are many evidences that people find it easier to deal with positive information than with negative: it is easier to decide the truth or falsity of positive than of negative sentences (Wason, 1959, 1961); the assertion that something is absent takes longer to comprehend than the assertion that something is present (Clark, 1974); and inferences from negative premises require more time to make or evaluate and are more likely to be erroneous or evaluated incorrectly than are those that are based on positive premises (Fodor, Fodor, & Garrett, 1975). How far the idea can be pushed is a question for research. I suspect that failure to try to construct counterarguments or to find counterevidence is a major and relatively pervasive weakness of human reasoning.

In any case, the positivity bias itself requires an explanation. Does such a bias have some functional value? Is it typically more important for people to be attuned to occurrences than to nonoccurrences of possible events? Is language processed more effectively by one who is predisposed to hear positive rather than negative assertions? Is it generally more important to be able to make valid inferences from positive than from negative premises? It would not be surprising to discover that a positivity bias is advantageous in certain ways, in which case Bacon's dictum that we should be indifferently disposed toward positives and negatives would be wrong.

There is also the possibility that the positivity bias is, at least in part, motivationally based. Perhaps it is the case that the processing of negative information generally takes more effort than does the processing of positive information

and often we are simply not willing to make the effort that adequate processing of the negative information requires. As to why the processing of negative information should require more effort than the processing of positive information, perhaps it is because positive information more often than not is provided by the situation, whereas negative information must be actively sought, from memory or some other source. It is generally clear from the statement of a hypothesis, for example, what a positive instance of the hypothesized event would be, whereas that which constitutes a negative instance may require some thought.

Conditional Reference Frames

Several investigators have shown that when people are asked to explain or imagine why a hypothesis might be true or why a possible event might occur, they tend to become more convinced that the hypothesis is true or that the event will occur, especially if they have not given much thought to the hypothesis or event before being asked to do so (Campbell & Fairey, 1985; Hirt & Sherman, 1985; Sherman et al., 1983). In some cases, people who were asked to explain why a particular event had occurred and then were informed that the event did not occur after they did so, still considered the event more "likely" than did others who were not asked to explain why it might occur (Ross et al., 1977).

Koehler (1991), who has reviewed much of the work on how explanations influence beliefs, has suggested that producing an explanation is not the critical factor and that simply coming up with a focal hypothesis is enough to increase one's confidence in it. Anything, he suggested, that induces one to accept the truth of a hypothesis temporarily will increase one's confidence that it is, in fact, true. Calling attention to a specified hypothesis results in the establishment of a focal hypothesis, and this in turn induces the adoption of a *conditional reference frame*, in which the focal hypothesis is assumed to be true.

Adoption of a conditional reference frame influences subsequent hypothesis-evaluation processes in three ways, Koehler (1991) suggested: it affects the way the problem is perceived, how relevant evidence is interpreted, and the direction and duration of information search. According to this author,

> Once a conditional reference frame has been induced by an explanation task, a certain inertia sets in, which makes it more difficult to consider alternative hypotheses impartially. In other words, the initial impression seems to persist despite the person's efforts to ignore it while trying to give fair consideration to an alternative view. (p. 503)

Hoch's (1984) findings regarding people who were asked to generate reasons for expecting a specified event and reasons against expecting it support Koehler's conclusion: in Hoch's study those who generated the pro reasons first and the con reasons later considered the event more probable than did those who generated the pro and con reasons in the opposite order. Koehler related the phenomenon to that of *mental set* or *fixedness* that is sometimes described in discussions of problem solving. He argued that adopting a conditional reference frame to determine confidence in a hypothesis is probably a good general method but that, like other heuristic methods, it also can yield overconfidence in some instances.

In a recent study of gambling behavior, Gibson, Sanbonmatsu, and Posavac (1997) found that participants who were asked to estimate the probability that a particular team would win the NBA basketball championship made higher estimates of that team's probability of winning than did control participants who were not asked to focus on a single team; the focused participants also were more willing to bet on the focal team. This suggests that focusing on one among several possible event outcomes, even as a consequence of being arbitrarily forced to do so, can have the effect of increasing the subjective likelihood of that outcome.

Pragmatism and Error Avoidance

Much of the discussion of confirmation bias is predicated on the assumption that, in the situations in which it has generally been observed, people have been interested in determining the truth or falsity of some hypothesis(es) under consideration. But determining its truth or falsity is not the only, or necessarily even the primary, objective one might have with respect to a hypothesis. Another possibility is that of guarding against the making of certain types of mistakes.

In many real-life situations involving the evaluation of a meaningful hypothesis, or the

making of a choice with an outcome that really matters to the one making it, some ways of being wrong are likely to be more regretable than others. Investigators have noted that this fact makes certain types of biases functional in specific situations (Cosmides, 1989; Friedrich, 1993; Hogarth, 1981; Schwartz, 1982). When, for example, the undesirable consequences of judging a true hypothesis to be false are greater than those of judging a false hypothesis to be true, a bias toward confirmation is dictated by some normative models of reasoning and by common sense.

Friedrich (1993) argued that "our inference processes are first and foremost pragmatic, survival mechanisms and only secondarily truth detection strategies" (p. 298). In this view, peoples' inferential strategies are well suited to the identification of potential rewards and the avoidance of costly errors, but not to the objective of hypothesis testing in accordance with the logic of science. Inference strategies that are often considered to be seriously flawed not only may have desired effects in real-world contexts, but may also be seen as correct when judged in terms of an appropriate standard.

To illustrate the point, Friedrich (1993) used the example of an employer who wants to test the hunch that extroverts make the best salespeople. If the employer checked the sales performance only of extroverts, found it to be very good and, on this basis, decided to hire only extroverts for sales positions, one would say that she had not made an adequate test of her hunch because she did not rule out the possibility that introverts might do well at sales also. But if her main objective was to avoid hiring people who will turn out to be poor at sales, satisfying herself that extroverts make good sales people suffices; the fact that she has not discovered that introverts can be good at sales too could mean that she will miss some opportunities by not hiring them, but it does not invalidate the decision to hire extroverts if the objective is to ensure that poor performers do not get hired.

Schwartz (1982) also argued that when responses have consequences that really matter, people are more likely to be concerned about producing desirable outcomes than about determining the truth or falsity of hypotheses. Contingent reinforcement may create functional behavioral units that people tend to repeat because the behaviors have worked in the past, and a bias toward confirmation is one of the stereotyped forms of behavior in which the operation of these units manifests itself. In performing a rule-discovery task, one may be attempting to maximize the probability of getting a positive response, which is tantamount to seeking a rule that is sufficient but not necessary to do so. Given that one has identified a condition that is sufficient for producing desired outcomes, there may be no compelling reason to attempt to determine whether that condition is also necessary. Schwartz pointed out too that, especially in social situations such as some of those studied by Snyder and colleagues attempting to evaluate hypotheses by falsification would require manipulating people and could involve considerable social cost.

Baron (1994) noted that truth seeking or hypothesis testing often may be combined with one or more other goals, and that one's behavior then also must be interpreted in the light of the other goal(s). If, for example, one is curious as to why a cake turned out well despite the fact that certain ingredients were substituted for those called for by the recipe, one may be motivated to explore the effects of the substitutions in such a way that the next experimental cake is likely to turn out well too.

When using a truth-seeking strategy would require taking a perceived risk, survival is likely to take precedence over truth finding, and it is hard to argue that rationality would dictate otherwise. It would seem odd to consider irrational the refusal, say, to eat mushrooms that one suspected of being poison because the decision is calculated to preserve one's wellbeing rather than to shed light on the question of whether the suspicion is indeed true. In general, the objective of avoiding disastrous errors may be more conducive to survival than is that of truth determination.

The desire to avoid a specific type of error may coincidentally dictate the same behavior as would the intention to determine the truth or falsity of a hypothesis. When this is the case, the behavior itself does not reveal whether the individual's intention is to avoid the error or to test the hypothesis. Friedrich (1993) suggested that some behavior that has been taken as evidence of people's preference for normative diagnostic tests of hypotheses and their interest in accuracy could have been motivated instead by the desire to avoid specific types of errors. In

other words, even when behavior is consistent with the assumption of truth seeking, it sometimes may be equally well interpreted, according to this view, in terms of error-minimizing strategies.

The assumption that decisions made or conclusions drawn in many real-life situations are motivated more by a desire to accomplish specific practical goals or to avoid certain types of errors than by the objective of determining the truth or falsity of hypotheses is a plausible one. Pragmatic considerations of this sort could often lead one to accept a hypothesis as true—to behave as though it were true—on less than compelling evidence that it is so, thus constituting a confirmation bias of sorts.

Educational Effects

At all levels of education, stress is placed on the importance of being able to justify what one believes. I do not mean to question the appropriateness of this stress, but I do want to note the possibility that, depending on how it is conveyed, it can strengthen a tendency to seek confirming evidence selectively or establish such a tendency if it does not already exist. If one is constantly urged to present reasons for opinions that one holds and is not encouraged also to articulate reasons that could be given against them, one is being trained to exercise a confirmation bias.

Narveson (1980) noted that when students write compositions, they typically evaluate their claims by considering supporting evidence only. He argued that standard methods for teaching composition foster this. The extent to which the educational process makes explicit the distinction between case-building and evidence-weighing deserves more attention. If the distinction is not made and what is actually case-building passes for the impartial use of evidence, this could go some way toward accounting for the pervasiveness and strength of the confirmation bias among educated adults.

Ideally, one would like students, and people in general, to evaluate evidence objectively and impartially in the formation and evaluation of hypotheses. If, however, there are fairly pervasive tendencies to seek or give undue weight to evidence that is confirmatory with respect to hypotheses that people already hold and to avoid or discount evidence that is disconfirmatory

with respect to them, there is a need to be especially sensitive to the educational practices that could serve to strengthen an already strong bias.

Utility of Confirmation Bias

Most commentators, by far, have seen the confirmation bias as a human failing, a tendency that is at once pervasive and irrational. It is not difficult to make a case for this position. The bias can contribute to delusions of many sorts, to the development and survival of superstitions, and to a variety of undesirable states of mind, including paranoia and depression. It can be exploited to great advantage by seers, soothsayers, fortune tellers, and indeed anyone with an inclination to press unsubstantiated claims. One can also imagine it playing a significant role in the perpetuation of animosities and strife between people with conflicting views of the world.

Even if one accepts the idea that the confirmation bias is rooted more in cognitive limitations than in motivation, can anyone doubt that whenever one finds oneself engaged in a verbal dispute it becomes very strong indeed? In the heat of an argument people are seldom motivated to consider objectively whatever evidence can be brought to bear on the issue under contention. One's aim is to win and the way to do that is to make the strongest possible case for one's own position while countering, discounting, or simply ignoring any evidence that might be brought against it. And what is true of one disputant is generally true of the other, which is why so few disputes are clearly won or lost. The more likely outcome is the claim of victory by each party and an accusation of recalcitrance on the part of one's opponent.

But whenever an apparently dysfunctional trait or behavior pattern is discovered to be pervasive, the question arises as to how, if it is really dysfunctional, did it get to be so widespread. By definition, dysfunctional tendencies should be more prone to extinction than functional ones. But aspects of reasoning that are viewed as flawed from one perspective sometimes can be considered appropriate, perhaps because they are adaptively useful in certain real-world situations from another perspective (Arkes, 1991; Funder, 1987; Greenwald, 1980). Does the confirmation bias have

some adaptive value? Does it serve some useful purpose(s)?

Utility in Science

According to the principle of *falsifiability* (Popper, 1959), an explanation (theory, model, hypothesis) cannot qualify as scientific unless it is falsifiable in principle. This is to say there must be a way to show the explanation to be false if in fact it is false. Popper focused on falsifiability as the distinguishing characteristic of the scientific method because of a conviction that certain theories of the day (in particular, Marx's theory of history, Freud's theory of psychoanalysis, and Adler's individual psychology) "appeared to be able to explain practically everything that happened within the fields to which they referred" (Popper, 1962/1981, p. 94). Thus subscribers to any one of these theories were likely to see confirming evidence anywhere they looked. According to Popper (1959), the most characteristic element in this situation seemed to be the incessant stream of confirmations, of observations which 'verified' the theories in question; and this point was constantly emphasized by their adherents" (p. 94). There was, in Popper's view, no conceivable evidence that could be brought to bear on any of these theories that would be viewed by an adherent as grounds for judging it to be false. Einstein's theory of relativity impressed Popper as being qualitatively different in the important respect that it made risky predictions, which is to say predictions that were incompatible with specific possible results of observation. This theory was, in principle, refutable by empirical means and therefore qualified as scientific according to Popper's criterion.

Although Popper articulated the principle of falsifiability more completely than anyone before him, the idea has many antecedents in philosophy and science. It is foreshadowed, for example, in the Socratic method of refutation (*elenchos*), according to which what is to be taken as truth is whatever survives relentless efforts at refutation (MacIntyre, 1988). Lakatos (1976, 1978) and Polya (1954a, 1954b) have discussed the importance of this attitude in reasoning in mathematics—especially in proof-making—at length. The principle of falsifiability was also anticipated by T. H. Huxley (1894/1908), who spoke of "a beautiful hypoth-

esis killed by an ugly fact," and by David Hartley (1748/1981), who proposed a *rule of false*. According to this rule, the acceptability of any supposition or hypothesis should be its ability to provide the basis for the deduction of observable phenomena: "He that forms hypotheses from the first, and tries them by the facts, soon rejects the most unlikely ones; and being freed from these, is better qualified for the examination of those that are probable" (p. 90).

Hypotheses are strengthened more when highly competent scientists make concerted efforts to disprove them and fail than when efforts at disproof are made by less competent investigators or in are made half-hearted ways. As Polya (1954a) put it, "the more danger, the more honor." What better support could Einstein's corpuscular theory of light have received than Millikan's failure to show it to be wrong, despite 10 years of experimentation aimed at doing so? (It is interesting to note that throughout this period of experimentation, Millikan continued to insist on the untenability of the theory despite his inability to show by experiment its predictions to be in error.)

Despite the acceptance of the falsifiability principle by the scientific community as a whole, one would look long and hard to find an example of a well-established theory that was discarded when the first bit of disconfirming evidence came to light. Typically an established theory has been discarded only after a better theory has been offered to replace it. Perhaps this should not be surprising. What astronomer, Waismann (1952) asked, would abandon Kepler's laws on the strength of a single observation? Scientists have not discarded the idea that light travels at a constant speed of about 300,000 kilometers per second and that nothing can travel faster simply because of the discovery of radio sources that seem to be emanating from a single quasar and moving away from each other at more than nine times that speed (Gardner, 1976). It appears that, the principle of falsifiability notwithstanding, "Science proceeds on preponderance of evidence, not on finality" (Drake, 1980, p. 55).

Application of the falsifiability principle to the work of individual scientists seems to indicate that when one comes up with a new hypothesis, one should immediately try to falsify it. Common sense suggests this too; if the hypothesis is false, the sooner one finds that out,

the less time one will waste entertaining it. In fact, as has already been noted, there is little evidence that scientists work this way. To the contrary, they often look much harder for evidence that is supportive of a hypothesis than for evidence that would show it to be false.

Kepler's laborious effort to find a connection between the perfect polyhedra and the planetary orbits is a striking example of a search by a scientist for evidence to confirm a favored hypothesis. Here is his account of the connection he finally worked out and his elation upon finding it, as quoted in Boorstin (1985):

> The earth's orbit is the measure of all things; circumscribe around it a dodecahedron, and the circle containing this will be Mars; circumscribe around Mars a tetrahedron, and the circle containing this will be Jupiter; circumscribe around Jupiter a cube, and the circle containing this will be Saturn. Now inscribe within the earth an icosahedron, and the circle contained in it will be Mercury. You now have the reason for the number of planets. . . . This was the occasion and success of my labors. And how intense was my pleasure from this discovery can never be expressed in words. I no longer regretted the time wasted. Day and night I was consumed by the computing, to see whether this idea would agree with the Copernican orbits, or if my joy would be carried away by the wind. Within a few days everything worked, and I watched as one body after another fit precisely into its place among the planets. (p. 310)

People are inclined to make light of this particular accomplishment of Kepler's today, but it was a remarkable intellectual feat. The energy with which he pursued what he saw as an intriguing clue to how the world works is inspiring. Bell (1946/1991) argued that it was Kepler's "Pythagorean faith in a numerical harmony of the universe" that sustained him "in his darkest hours of poverty, domestic tragedy, persecution, and twenty-two years of discouragement as he calculated, calculated, calculated to discover the laws of planetary orbits" (p. 181). The same commitment that kept Kepler in pursuit of confirmation of his polyhedral model of planetary orbits yielded the three exquisitly beautiful—and correct—laws of planetary motion for which he is honored today.

My point is not to defend the confirmation bias as an effective guide to truth or even as a heuristically practical principle of logical thinking. It is simply to note that the quest to find support for a particular idea is a common phenomenon in science (I. B. Cohen, 1985; Holton, 1973), that such a quest has often

provided the motivation to keep scientists working on demanding intellectual problems against considerable odds, and that the resulting work has sometimes yielded lasting, if unexpected, results.

Students of the scientific process have noted the conservatism of science as an institution (I. B. Cohen, 1985; T. S. Kuhn, 1970), and illustrations of it were given in an earlier part of this article. This conservatism can be seen as an institutional confirmation bias of sorts. Should we view such a bias as beneficial overall or detrimental to the enterprise? An extensive discussion of this question is beyond the scope of this article. However, it can be argued that a degree of conservativism plays a stabilizing role in science and guards the field against uncritical acceptance of so-called discoveries that fail to stand the test of time.

Price (1963) referred to conservatism in the body of science as "a natural counterpart to the open-minded creativity that floods it with too many new ideas" (p. 64). Justification for a certain degree of conservatism is found in the embarrassment that the scientific community has occasionally experienced as a consequence of not being sufficiently sceptical of new discoveries. The discovery of magnetic monopoles, which was widely publicized before a close examination of the evidence forced more guarded interpretations, and that of polywater, which motivated hundreds of research projects over decade following its discovery in the 1960s, are examples.

The scientific community's peremtory rejection of Wegener's (1915/1966) theory of continental drift when it was first put forth is often held out as an especially egregious example of excessive—and self-serving—conservatism on the part of scientists. The view has also been expressed, however, that the geologists who dismissed Wegener's theory, whatever their motivations, acted in a way that was conducive to scientific success. Solomon (1992), who took this position, acknowledged that the behavior was biased and motivated by the desire to protect existing beliefs, but she argued that "bias and belief perseverance made possible the distribution of effort, and this in turn led to the advancement of the debate over [continental] drift" (p. 443).

Solomon's (1992) review of this chapter in the history of geological research makes it clear

that the situation was not quite as simple as some accounts that focus on the closed-mindedness of the geologists at the time would lead one to believe. The idea of continental drift was not entirely new with Wegener, for example, although Wegener was the first to propose a well-developed theory. A serious limitation of the theory was its failure to identify a force of sufficient magnitude to account for the hypothesized movement of continents. (The notion of plate tectonics and evidence regarding sea-floor spreading came much later; LeGrand, 1988.)

Solomon (1992) argued that it was, in part, because Wegener was not trained as a geologist and therefore not steeped in the "stabilist" theories of the time, that his thinking, relatively unconstrained by prior beliefs about the stability of continents, could easily embrace a possibility that was so contrary to the prevailing view. She pointed out too that when geologists began to accept the notion of drift, as evidence favoring it accumulated, it was those with low publication rates who were the most likely to do so: "their beliefs were less entrenched (cognitively speaking) than those who had reasoned more and produced more, so belief perseverance was less of an impediment to acceptance of drift" (p. 449).

Although Solomon (1992) argued that bias and belief perseverance were responsible for much of the distribution of research effort that led finally to the general acceptance of the theory of continental drift, and to that of plate tectonics to which it led in turn, she does not claim that a productive distribution could not have been effected without the operation of these factors. The question of whether these factors facilitated or impeded progress in geology remains an unanswered one; it is not inconceivable that progress could have been faster if the distribution of effort were determined on some other basis.

In any case, it can be argued that a certain degree of conservativism serves a useful stabilizing role in science and is consistent with, if not dictated by, the importance science attaches to testability and empirical validation. Moreover, if Ziman (1978) is right, the vast majority of the new hypotheses put forward by scientists prove to be wrong:

> Even in physics, there is no infallible procedure for generating reliable knowledge. The calm order and perfection of well-established theories, accredited by innumerable items of evidence from a thousand different hands, eyes and brains, is not characteristic of the front-line of research, where controversy, conjecture, contradiction and confusion are rife. The physics of undergraduate test-books is 90% true; the contents of the primary research journals of physics is 90% false. (p. 40)

"According to termperament," Ziman noted, "one may be impressed by the coherence of well-established theories, or horrified by the contradictions of knowledge in the making" (p. 100).

Fischhoff and Beyth-Marom (1983) made a related point when commenting on Mahoney's (1977) finding that scientists tended to be less critical of a fictitious study that reported results supportive of the dominant hypothesis in their field than of one that reported results that were inconsistent with it. They noted that a reluctance by scientists to relinquish pet beliefs is only one interpretation that could be put on the finding. Another possibility is that what appears to be biased behavior reflects "a belief that investigators who report disconfirming results tend to use inferior research methods (e.g., small samples leading to more spurious results), to commit common mistakes in experimental design, or, simply, to be charlatans" (Fischoff and Beyth-Marom, 1983, p. 251). To the extent that such a belief is accurate, a bias against the ready acceptance of results that are disconfirming of prevailing hypotheses can be seen as a safeguard against precipitous changes of view that may prove to be unjustified. It does not follow, of course, that this is the only reason for a confirmation bias in science or the only effect.

Is Belief Perseverance Always Bad?

It is easy to see both how the confirmation bias helps preserve existing beliefs, whether true or false, and how the perseverance of unjustified beliefs can cause serious problems. Is there anything favorable to be said about a bias that tends to perpetuate beliefs independently of their factuality? Perhaps, at least from the narrow perspective of an individual's mental health. It may help, for example, to protect one's ego by making one's favored beliefs less vulnerable than they otherwise would be. Indeed, it seems likely that a major reason why the confirmation bias is so ubiquitous and so enduring is its effectiveness in preserving

preferred beliefs and opinions (Greenwald, 1980).

But even among people who might see some benefit to the individual in a confirmation bias, probably few would contest the claim that when the tendency to persevere in a belief is so strong that one refuses to consider evidence that does not support that belief, it is irrational and offends our sense of intellectual honesty. That is not to say that dogmatic confidence in one's own beliefs and intolerance of opposing views can never work to one's advantage. Boorstin (1958) argued, for example, that it was precisely these qualities that permitted the 17th-century New England Puritans to establish a society with the ingredients necessary for survival and prosperity. He wrote,

> Had they spent as much of their energy in debating with each other as did their English contemporaries, they might have lacked the single-mindedness needed to overcome the dark, unpredictable perils of a wilderness. They might have merited praise as precursors of modern liberalism, but they might never have helped found a nation. (p. 9)

Contrary to the popular stereotype of the Puritans, they were not preoccupied with religious dogma but rather with more practical matters because, as Boorstin noted, they had no doubts and allowed no dissent. They worried about such problems as how to select leaders and representatives, the way to establish the proper limits of political power, and how to construct a feasible federal organization.

The question of the conditions under which one should retain, reject, or modify an existing belief is a controversial one (Cherniak, 1986; Harman, 1986; Lycan, 1988). The controversy is not likely to be settled soon. Whatever the answer to the question is, the confirmation bias must be recognized as a major force that works against easy and frequent opinion change. Probably very few people would be willing to give up long-held and valued beliefs on the first bit of contrary evidence found. It is natural to be biased in favor of one's established beliefs. Whether it is rational is a complicated issue that can too easily be treated simplistically; however, the view that a person should be sufficiently objective and open minded to be willing to toss out any belief upon the first bit of evidence that it is false seems to me wrong for several reasons.

Many, perhaps most, of the beliefs that matter to individuals tend not to be the type that can be falsified, in the Popperian sense, by a single counterindicative bit of data. They tend rather to be beliefs for which both supportive and counterindicative evidence can be found, and the decision as to whether to hold them is appropriately made on the basis of the relative weights or merits of the pro and con arguments.

Second, it is possible to hold a belief for good and valid reasons without being able to produce all of those reasons on demand. Some beliefs are shaped over many years, and the fact that one cannot articulate every reason one has or has ever had for a particular one of them does not mean that it is unfounded. Also, as Nisbett and Ross (1980) pointed out, there are practical time constraints that often limit the amount of processing of new information one can do. In view of these limitations, the tendency to persevere may be a stabilizing hedge against overly frequent changes of view that would result if one were obliged to hold only beliefs that one could justify explicitly at a moment's notice. This argument is not unlike the one advanced by Blackstone (1769/1962) in defense of not lightly scuttling legal traditions.

Third, for assertions of the type that represent basic beliefs, there often are two ways to be wrong: to believe false ones, or to disbelieve true ones. For many beliefs that people hold, these two possibilities are not equally acceptable, which is to say that an individual might consider it more important to avoid one type of error than the other. This is, of course, the argument behind Pascal's famous wager.

To argue that it is not necessarily irrational to refuse to abandon a long-held belief upon encountering some evidence that appears contradictory is not to deny that there is such a thing as holding on to cherished beliefs too tenaciously and refusing to give a fair consideration to counterindicative evidence. The line between understandable conservativism with respect to changing established beliefs and obstinate closed-mindedness is not an easy one to draw. But clearly, people sometimes persevere beyond reason. Findings such as those of Pitz (1969), Pitz et al. (1967), Lord et al. (1979), and especially Ross et al. (1975), who showed that people sometimes persevere in beliefs even when the evidence on which the beliefs were initially formed has been demonstrated to them to be fraudulent, have provided strong evidence of that fact.

Confirmation Bias Compounding Inadequate Search

The idea that inadequate effort is a basic cause of faulty reasoning is common among psychologists. Kanouse (1972) suggested that people may be satisfied to have an explanation for an event that is sufficient and not feel the need to seek the best of all possibilities. Nisbett and Ross (1980) supported the same idea and suggested that it is especially the case when attempting to come up with a causal explanation:

> The lay scientist seems to search only until a plausible antecedent is discovered that can be linked to the outcome through some theory in the repertoire. Given the richness and diversity of that repertoire, such a search generally will be concluded quickly and easily. A kind of vicious cycle results. The subjective ease of explanation encourages confidence, and confidence makes the lay scientist stop searching as soon as a plausible explanation is adduced, so that the complexities of the task, and the possibilities for "alternative explanations" no less plausible than the first, are never allowed to shake the lay scientist's confidence. (p. 119, 120)

Perkins and his colleagues expressed essentially the same idea with their characterization of people as *make-sense epistimologists* (Perkins et al., 1983, 1991). The idea here is that people think about a situation only to the extent necessary to make sense—perhaps superficial sense—of it:

> When sense is achieved, there is no need to continue. Indeed, because further examination of an issue might produce contrary evidence and diminish or cloud the sense of one's first pass, there is probably reinforcement for early closure to reduce the possibility of cognitive dissonance. Such a makes-sense approach is quick, easy, and, for many purposes, perfectly adequate. (Perkins et al., 1991, p. 99)

Baron (1985, 1994) and Pyszczynski and Greenberg (1987) also emphasized insufficient search as the primary reason for the premature drawing of conclusions.

All of these characterizations of the tendency of people to do a less-than-thorough search through the possibilities before drawing conclusions or settling on causal explanations are consistent with Simon's (1957, 1983/1990) view of humans as satisficers, as opposed to optimizers or maximizers. The question of interest in the present context is that of how the criterion for being satisfied should be set. Given that search is seldom exhaustive and assuming that,

in many cases, it cannot be, how much should be enough? How should one decide when to stop?

Without at least tentative answers to these types of questions, it is difficult to say whether any particular stopping rule should be considered rational. Despite this vagueness with respect to criteria, I believe that the prevailing opinion among investigators of reasoning is that people often stop—come to conclusions, adopt explanations—before they should, which is not to deny that they may terminate a search more quickly when time is at a premium and search longer when accuracy is critical (Kruglanski, 1980; Kruglanski & Ajzen, 1983; Kruglanski & Freund, 1983). The search seems to be not only less than extensive but, in many cases, minimal, stopping at the first plausible endpoint.

If this view is correct, the operation of a confirmation bias will compound the problem. Having once arrived at a conclusion, belief, or point of view, however prematurely, one may thereafter seek evidence to support that position and interpret newly acquired information in a way that is partial to it, thereby strengthening it. Instead of making an effort to test an initial hypothesis against whatever counterindicative evidence might be marshalled against it, one may selectively focus on what can be said in its favor. As the evidence favoring the hypothesis mounts, as it is bound to do if one gives credence to evidence that is favorable and ignores or discounts that that is not, one will become increasingly convinced of the correctness of the belief one originally formed.

Concluding Comments

We are sometimes admonished to be tolerant of the beliefs or opinions of others and critical of our own. Laplace (1814/1956), for example, gave this eloquent advice:

> What indulgence should we not have . . . for opinions different from ours, when this difference often depends only upon the various points of view where circumstances have placed us! Let us enlighten those whom we judge insufficiently instructed; but first let us examine critically our own opinions and weigh with impartiality their respective probabilities. (p. 1328)

But can we assess the merits of our own opinions impartially? Is it possible to put a belief that one holds in the balance with an opposing belief that one does not hold and give them a fair weighing? I doubt that it is. But that

is not to say that we cannot hope to learn to do better than we typically do in this regard.

In the aggregate, the evidence seems to me fairly compelling that people do not naturally adopt a falsifying strategy of hypothesis testing. Our natural tendency seems to be to look for evidence that is directly supportive of hypotheses we favor and even, in some instances, of those we are entertaining but about which are indifferent. We may look for evidence that is embarrassing to hypotheses we disbelieve or especially dislike, but this can be seen as looking for evidence that is supportive of the complementary hypotheses. The point is that we seldom seem to seek evidence naturally that would show a hypothesis to be wrong and to do so because we understand this to be an effective way to show it to be right if it really is right.

The question of the extent to which the confirmation bias can be modified by training deserves more research than it has received. Inasmuch as a critical step in dealing with any type of bias is recognizing its existence, perhaps simply being aware of the confirmation bias—of its pervasiveness and of the many guises in which it appears—might help one both to be a little cautious about making up one's mind quickly on important issues and to be somewhat more open to opinions that differ from one's own than one might otherwise be.

Understanding that people have a tendency to overestimate the probable accuracy of their judgments and that this tendency is due, at least in part, to a failure to consider reasons why these judgments might be inaccurate provides a rationale for attempting to think of reasons for and (especially) against a judgment that is to be made. Evidence that the appropriateness of people's confidence in their judgments can be improved as a consequence of such efforts is encouraging (Arkes, Faust, Guilmette, & Hart, 1988; Hoch, 1984, 1985; Koriat et al., 1980).

The knowledge that people typically consider only one hypothesis at a time and often make the assumption at the outset that that hypothesis is true leads to the conjecture that reasoning might be improved by training people to think of alternative hypotheses early in the hypothesis-evaluation process. They could be encouraged to attempt to identify reasons why the complement of the hypothesis in hand might be true. Again, the evidence provides reason for optimism that the approach can work (C. A.

Anderson, 1982; C. A. Anderson & Sechler, 1986; Lord, Lepper, & Preston, 1984).

To the extent that what appear to be biases are sometimes the results of efforts to avoid certain types of decision errors (Friedrich, 1993), making these other types of possible errors more salient may have a debiasing effect. On the other hand, if the errors that one is trying to avoid are, in fact, more costly than bias errors, such debiasing might not be desirable in all instances. Here the distinction between the objective of determining the truth or falsity of a hypothesis and that of avoiding an undesirable error (at the expense of accepting the possibility of committing a less undesirable one) is an important one to keep in mind.

Finally, I have argued that the confirmation bias is pervasive and strong and have reviewed evidence that I believe supports this claim. The possibility will surely occur to the thoughtful reader that what I have done is itself an illustration of the confirmation bias at work. I can hardly rule the possibility out; to do so would be to deny the validity of what I am claiming to be a general rule.

References

Alloy, L. B., & Abramson, L. Y. (1980). The cognitive component of human helplessness and depression: A critical analysis. In J. Garber & M. E. P. Seligman (Eds.), *Human helplessness: Theory and applications*. New York: Academic Press.

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,* 112–148.

Anderson, C. A. (1982). Inoculation and counterexplanation: Debiasing techniques in the perseverance of social theories. *Social Cognition, 1,* 126–139.

Anderson, C. A., & Sechler, E. S. (1986). Effects of explanation and counterexplanation on the development and use of social theories. *Journal of Personality and Social Psychology, 50,* 24–34.

Anderson, N. H., & Jacobson, A. (1965). Effect of stimulus inconsistency and discounting instructions in personality impression formation. *Journal of Personality and Social Psychology, 2,* 531–539.

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin, 110,* 486–498.

Arkes, H. R., Dawes, R. M., & Christensen, C. (1986). Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes, 37,* 93–110.

Arkes, H. R., Faust, D., Guilmette, T. J., & Hart, K. (1988). Elimination of the hindsight bias. *Journal of Applied Psychology, 73*, 305–307.

Bacon, F. (1939). *Novum organum.* In Burtt, E. A. (Ed.), *The English philosophers from Bacon to Mill* (pp. 24–123). New York: Random House. (Original work published in 1620)

Baddeley, A. D., & Woodhead, M. (1983). Improving face recognition ability. In S. M. A. Lloyd-Bostock & B. R. Clifford (Eds.), *Evaluating witness evidence.* Chichester, England: Wiley.

Barber, B. (1961). Resistence by scientists to scientific discovery. *Science, 134*, 596–602.

Baron, J. (1985). *Rationality and intelligence.* New York: Cambridge University Press.

Baron, J. (1991). Beliefs about thinking. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 169–186). Hillsdale, NJ: Erlbaum.

Baron, J. (1994). *Thinking and deciding* (2nd ed.). New York: Cambridge University Press.

Baron, J. (1995). Myside bias in thinking about abortion. *Thinking and reasoning, 7*, 221–235.

Baron, J., Beattie, J., & Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning II: Congruence, information, and certainty. *Organizational Behavior and Human Decision Processes, 42*, 88–110.

Barrows, H. S., Feightner, J. W., Neufeld, V. R., & Norman, G. R. (1978). *Analysis of the clinical methods of medical students and physicians.* Hamilton, Ontario, Canada: McMaster University School of Medicine.

Barrows, H. S., Norman, G. R., Neufeld, V. R., & Feightner, J. W. (1977). Studies of the clinical reasoning process of medical students and physicians. *Proceedings of the sixteenth annual conference on research in medical education.* Washington, DC: Association of American Medical Colleges.

Bassock, M., & Trope, Y. (1984). People's strategies for testing hypotheses about another's personality: Confirmatory or diagnostic? *Social Cognition, 2*, 199–216.

Beattie, J., & Baron, J. (1988). Confirmation and matching biases in hypothesis testing. *Quarterly Journal of Experimental Psychology, 40A*, 269–298.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders.* New York: International Universities Press.

Bell, E. T. (1991). *The magic of numbers.* New York: Dover. (Original work published 1946)

Berwick, D. M., Fineberg, H. V., & Weinstein, M. C. (1981). When doctors meet numbers. *American Journal of Medicine, 71*, 991.

Beyth-Marom, R., & Fischhoff, B. (1983). Diagnosticity and pseudodiagnosticity. *Journal of Personality and Social Research, 45*, 1185–1197.

Blackstone, W. (1962). *Commentaries on the laws of England of public wrongs.* Boston: Beacon. (Original work published 1769)

Boorstin, D. J. (1958). *The Americans: The colonial experience.* New York: Vintage Books.

Boorstin, D. J. (1985). *The discoverers: A history of man's search to know his world and himself.* New York: Vintage Books.

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

Bruner, J. S., & Potter, M. C. (1964). Interference in visual recognition. *Science, 144*, 424–425.

Camerer, C. (1988). Illusory correlations in perceptions and predictions of organizational traits. *Journal of Behavioral Decision Making, 1*, 77–94.

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.

Cassells, W., Schoenberger, A., & Grayboys, T. B. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine, 299*, 999.

Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal Learning and Verbal Behavior, 6*, 151–155.

Chapman, L. J., & Chapman, J. P. (1959). Atmosphere effect reexamined. *Journal of Experimental Psychology, 58*, 220–226.

Chapman, L. J., & Chapman, J. P. (1967a). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology, 72*, 193–204.

Chapman, L. J., & Chapman, J. P. (1967b). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology, 74*, 271–280.

Chapman, L. J. & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology, 74*, 271–280.

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

Chernev, A. (1997). The effect of common features on brand choice: Moderating the effect of attribute importance. *Journal of Consumer Research, 23*, 304–311.

Cherniak, C. (1986). *Minimal rationality.* Cambridge, MA: MIT Press.

Christensen-Szalanski, J. J., & Bushyhead, J. B. (1988). Physicians' use of probabilistic information in a real clinical setting. In J. Dowie & A. Elstein (Eds.), *Professional judgment. A reader in clinical decision making* (pp. 360–373). Cambridge, England: Cambridge University Press. (Original work published 1981)

Clark, H. H. (1974). Semantics and comprehension. In T. A. Sebeok (Ed.), *Current trends in linguistics.*

Volume 12: Linguistics and adjacent arts and sciences. The Hague, Netherlands: Mouton.

Clark, H. H., & Chase, W. B. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3,* 472–517.

Cohen, I. B. (1985). *Revolution in science.* Cambridge, MA: Harvard University Press.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences, 4,* 317–331.

Collins, S., & Pinch, J. (1993). *The Golem: What everyone should know about science.* New York: Cambridge University Press.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31,* 187–276.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 162–228). New York: Oxford University Press.

Crocker, J. (1981). Judgment of covariation by social perceivers. *Psychological Bulletin, 90,* 272–292.

Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist, 35,* 867–881.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labelling effects. *Journal of Personality and Social Psychology, 44,* 20–33.

DeSmet, A. A., Fryback, D. G., & Thornbury, J. R. (1979). A second look at the utility of radiographic skull examinations for trauma. *American Journal of Roentgenology, 132,* 95.

Devine, P. G., Hirt, E. R., & Gehrke, E. M. (1990). Diagnostic and confirmation strategies in trait hypothesis testing. *Journal of Personality and Social Psychology, 58,* 952–963.

Devine, P. G., & Ostrom, T. M. (1985). Cognitive mediation of inconsistency discounting. *Journal of Personality and Social Psychology, 49,* 5–21.

Doherty, M. E., & Mynatt, C. R. (1986). The magical number one. In D. Moates & R. Butrick (Eds.), *Inference Ohio University Interdisciplinary Conference 86* (Proceedings of the Interdisciplinary Conference on Inference; pp. 221–230). Athens: Ohio University.

Doherty, M. E., Mynatt, C. R., Tweney, R. D., & Schiavo, M. D. (1979). Pseudodiagnosticity. *Acta Psychologica, 43,* 111–121.

Drake, S. (1980). *Galileo.* New York: Oxford University Press.

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology, 34,* 590–598.

Dusek, J. B. (1975). Do teachers bias children's learning? *Review of Educational Research, 45,* 661–684.

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review, 85,* 395–416.

Elstein, A. S., & Bordage, G. (1979). Psychology of clinical reasoning. In G. Stone, F. Cohen, & N. Adler (Eds.) *Health psychology: A handbook.* San Francisco: Jossey-Bass.

Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). *Medical problem solving: An analysis of clinical reasoning.* Cambridge, MA: Harvard University Press.

Evans, J. St. B. T. (1972). On the problems of interpreting reasoning data: Logical and psychological approaches. *Cognition, 1,* 373–384.

Evans, J. St. B. T. (1982). *The psychology of deductive reasoning.* London: Routledge & Kegan Paul.

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

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

Evans, J. St., B. P., Newstead, E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction.* Hove, England: Erlbaum.

Farley, J., & Geison, G. L. (1974). Science politics and spontaneous generation in nineteenth-century France: The Pasteur-Pouchet debate. *Bulletin for the History of Medicine, 48,* 161–198.

Feldman, J. M., Camburn, A., & Gatti, G. M. (1986). Shared distinctiveness as a source of illusory correlation in performance appraisal. *Organizational Behavior and Human Decision Process, 37,* 34–59.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Fischhoff, B. (1977). Perceived informativeness of facts. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 349–358.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). Cambridge, England: Cambridge University Press.

Fischhoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review, 90,* 239–260.

Fodor, J. D., Fodor, J. A., & Garrett, M. F. (1975). The psychological unreality of semantic representations. *Linguistic Inquiry, 4,* 515–531.

Forer, B. (1949). The fallacy of personal validation: A classroom demonstration of gullibility. *Journal of Abnormal and Social Psychology, 44,* 118–123.

Foster, G., Schmidt, C., & Sabatino, D. (1976). Teacher expectancies and the label "learning disabilities." *Journal of Learning Disabilities, 9,* 111–114.

Freedman, J. L. (1964). Involvement, discrepancy, and change. *Journal of Abnormal and Social Psychology, 69,* 290–295.

Frey, D. (1986). Recent research on selective exposure to information. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 41–80). New York: Academic Press.

Friedrich, J. (1993). Primary error detection and minimization (PEDMIN) strategies in social cognition: A reinterpretation of confirmation bias phenomena. *Psychological Review, 100,* 298–319.

Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. *Psychological Bulletin, 101,* 79–90.

Gardner, M. (1957). *Fads and fallacies in the name of science.* New York: Dover.

Gardner, M. (1976). *The relativity explosion.* New York: Vintage Books.

Gibson, B., Sanbonmatsu, D. M., & Posavac, S. S. (1997). The effects of selective hypothesis testing on gambling. *Journal of Experimental Psychology: Applied, 3,* 126–142.

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

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition, 43,* 127–171.

Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Krüger, L. (1989). *The empire of chance: How probability changed science and everyday life.* New York: Cambridge University Press.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46,* 107–119.

Gilovich, T. (1983). Biased evaluation and persistence in gambling. *Journal of Personality and Social Psychology, 44,* 1110–1126.

Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life.* New York: Free Press.

Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory and Cognition, 10,* 597–602.

Goldberg, L. R. (1968). Simple models or simple processes? Some research in clinical judgment. *American Psychologist, 23,* 483–496.

Golding, S. L., & Rorer, L. G. (1972). Illusory correlation and subjective judgement. *Journal of Abnormal Psychology, 80,* 249–260.

Goodman, N. (1966). *The structure of appearance* (2nd ed.). Indianapolis, IN: Bobbs-Merrill.

Graesser, A. C., & Hemphill, D. (1991). Question answering in the context of scientific mechanisms. *Journal of Memory and Language, 30,* 186–209.

Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist, 35,* 603–618.

Griffin, D. W., Dunning, D., & Ross, L. (1990). The role of construal processes in overconfident predic-tions about the self and others. *Journal of Personality and Social Psychology, 59,* 1128–1139.

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

Griggs, R. A., & Cox, J. R. (1993). Permission schemas and the selection task. In J. St. B. T. Evans (Ed.), *The cognitive psychology of reasoning* (pp. 637–651). Hillsdale, NJ: Erlbaum.

Hamilton, D. L. (1979). A cognitive attributional analysis of stereotyping. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology,* (Vol. 12). New York: Academic Press.

Hamilton, D. L., Dugan, P. M., & Trolier, T. K. (1985). The formation of stereotypic beliefs: Further evidence for distinctiveness-based illusory correlations. *Journal of Personality and Social Psychology, 48,* 5–17.

Harman, G. (1986). *Change in view: Principles of reasoning.* Cambridge, MA: MIT Press.

Hartley, D. (1981). Hypotheses and the "rule of the false." In R. D. Tweney, M. E. Doherty, & C. R. Mynatt (Eds.), *Scientific thinking* (pp. 89–91). New York: Columbia University Press. (Original work published 1748)

Hasdorf, A. H., & Cantril, H. (1954). They saw a game. *Journal of Abnormal and Social Psychology, 49,* 129–134.

Hausch, D. B., Ziemba, W. T., & Rubenstein, M. (1981). Efficiency of the market for racetrack betting. *Management Science, 27,* 1435–1452.

Hawking, S. W. (1988). *A brief history of time: From the big bang to black holes.* New York: Bantam Books.

Hayden, T., & Mischel, W. (1976). Maintaining trait consistency in the resolution of behavioral inconsistency: The wolf in sheep's clothing? *Journal of Personality, 44,* 109–132.

Hempel, C. (1945). Studies in the logic of confirmation. *Mind, 54*(213), 1–26.

Hendry, S. H., & Shaffer, D. R. (1989). On testifying in one's own behalf: Interactive effects of evidential strength and defendant's testimonial demeanor on jurors' decisions. *Journal of Applied Psychology, 74,* 539–545.

Henle, M. (1962). On the relation between logic and thinking. *Psychological Review, 69,* 366–378.

Henrion, M., & Fischhoff, B. (1986). Assessing uncertainty in physical constants. *American Journal of Physics, 54,* 791–798.

Hirt, E. R., & Sherman, S. J. (1985). The role of prior knowledge in explaining hypothetical events. *Journal of Experimental Social Psychology, 21,* 519–543.

Hoch, S. J. (1984). Availability and inference in predictive judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 649–662.

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

Hoch, S. J., & Tschirgi, J. E. (1983). Cue redundancy and extra logical inferences in a deductive reasoning task. *Memory and Cognition, 11,* 200–209.

Hodgins, H. S., & Zuckerman, M. (1993). Beyond selecting information: Baises in spontaneous questions and resultant conclusions. *Journal of Experimental Social Psychology, 29,* 387–407.

Hogarth, R. M. (1981). Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. *Psychological Bulletin, 90,* 197–217.

Hollis, M. (1970). Reason and ritual. In B. R. Wilson (Ed.), *Rationality* (pp. 221–239). Oxford, England: Blackwell.

Holstein, J. A. (1985). Jurors' interpretations and jury decision making. *Law and Human Behavior, 9,* 83–99.

Holton, G. (1973). *Thematic origins of scientific thought.* Cambridge, MA: Harvard University Press.

Huxley, T. H. (1908). Discourses, biological and geological. In *Collected essays.* (Vol. 8). London: MacMillan and Company. (Original work published 1894)

Hyman, R. (1977). Cold reading. *The Skeptical Inquirer, 1,* 18–37.

Jennings, D., Amabile, T. M., & Ross, L. (1982). Informal covariation assessment: Data-based versus theory-based judgments. In A. Tversky, D. Kahneman, & P. Slovic (Eds.), *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

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

Jones, E. E., & Goethals, G. (1972). Order effects in impression formation: Attribution context and the nature of the entity. In E. E. Jones and others (Eds.), *Attribution: Perceiving the causes of behavior.* Morristown, NJ: General Learning Press.

Jones, E. E., Rock, L., Shaver, K. G., Goethals, G. R., & Ward, L. M. (1968). Pattern of performance and ability attribution: An unexpected primacy effect. *Journal of Personality and Social Psychology, 10,* 317–340.

Juslin, P. (1993). An explanation of the hard-easy effect in studies of realism of confidence in one's general knowledge. *European Journal of Cognitive Psychology, 5,* 55–71.

Juslin, P. (1994). The overconfidence phenomenon as a consequence of informal experimenter-guided selection of almanac items. *Organizational Behavior and Human Decision Processes, 57,* 226–246.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.

Kalven, H., & Zeisel, H. (1966). *The American jury.* Boston: Little, Brown.

Kanouse, D. E. (1972). Language, labeling, and attribution. In E. E. Jones, and others (Eds.), *Attribution: Perceiving the causes of behavior.* Morristown, NJ: General Learning Press.

Kelley, H. H. (1950). The warm-cold variable in first impressions of persons. *Journal of Personality, 18,* 431–439.

Keren, G. B. (1987). Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes, 39,* 98–114.

Kern, L., & Doherty, M. E. (1982). "Pseudodiagnosticity" in an idealized medical problem-solving environment. *Journal of Medical Education, 57,* 100–104.

Kidd, J. B. (1970). The utilization of subjective probabilities in production planning. *Acta Psychologica, 34,* 338–347.

Kirby, K. N. (1994a). False alarm: A reply to Over and Evans. *Cognition, 52,* 245–250.

Kirby, K. N. (1994b). Probabilities and utilities of fictional outcomes in Wason's four-card selection task. *Cognition, 51,* 1–28.

Klapper, J. T. (1960). *The effects of mass communications.* Glencoe, IL: Free Press.

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

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

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 107–118.

Kroger, J. K., Cheng, P. W., & Holyoak, K. J. (1993). Evoking the permission schema: The impact of explicit negation and a violation-checking context. In J. St. B. T. Evans (Ed.), *The cognitive psychology of reasoning* (pp. 615–635). Hillsdale, NJ: Erlbaum.

Kruglanski, A. W. (1980). Lay epistemology process and contents. *Psychological Review, 87,* 70–87.

Kruglanski, A. W., & Ajzen, I. (1983). Bias and error in human judgment. *European Journal of Social Psychology, 13,* 1–44.

Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology, 19,* 448–468.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review, 96,* 674–689.

Kuhn, D., Weinstock, M., & Flaton, R. (1994). How

well do jurors reason? Competence dimensions of individual variations in a juror reasoning task. *Psychological Science, 5,* 289–296.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed). Chicago: University of Chicago Press.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108,* 480–498.

Kunda, Z., & Nisbett, R. E. (1986). The psychometrics of everyday life. *Cognitive Psychology, 18,* 195–224.

Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery* (J. Worrall & E. Zahar, Eds.). New York: Cambridge University Press.

Lakatos, I. (1978). Falsification and the methodology of scientific research programmes. In I. Lakatos *The methodology of scientific research programmes* (J. Worrall & G. Currie, Eds.; pp. 8–101). New York: Cambridge University Press.

Langer, E. J., & Abelson, R. P. (1974). A patient by any other name. . . : Clinical group differences in labeling bias. *Journal of Consulting and Clinical Psychology, 42,* 4–9.

Laplace, P. S. (1956). Concerning probability. In J. R. Newman (Ed.), *The world of mathematics* (Vol. 2; pp. 1325–1333). New York: Simon and Schuster. (Original work published 1814)

Lawson, R. (1968). Order of presentation as a factor in jury persuasion. *Kentucky Law Journal, 56,* 523–555.

Lefford, A. (1946). The influence of emotional subject matter on logical reasoning. *Journal of General Psychology, 34,* 127–151.

LeGrand, H. E. (1988). *Drifting continents and shifting theories: The modern revolution in geology and scientific change.* Cambridge, England: Cambridge University Press.

Legrenzi, P. (1970). Relations between language and reasoning about deductive rules. In G. B. Flores d'Arcais & W. J. M. Levelt (Eds.), *Advances in psychologistics* (pp. 322–333). Amsterdam: North Holland.

Lenski, G. E., & Leggett, J. C. (1960). Caste, class, and deference in the research interview. *American Journal of Sociology, 65,* 463–467.

Levine, M. (1970). Human discrimination learning. The subset-sampling assumption. *Psychological Bulletin, 74,* 397–404.

Liberman, N., & Klar, Y. (1996). Hypothesis testing in Wason's selection task: Social exchange cheating detection or task understanding. *Cognition, 58,* 127–156.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance, 20,* 159–183.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1977). Calibration of probabilities: The state of the art. In H. Jungermann & G. de Zeeuw (Eds.), *Decision making and change in human affairs* (pp. 275–324). Dordrecht, Netherlands: Reidel.

Lingle, J. H., & Ostrom, T. M. (1981). Principles of memory and cognition in attitude formation. In R. E. Petty, T. M. Ostrom, & T. C. Brock (Eds.), *Cognitive responses in persuasive communications: A text in attitude change* (pp. 399–420). Hillsdale, NJ: Erlbaum.

Loftus, E. F., & Wagenaar, W. A. (1988). Lawyers' predictions of success. *Jurimetrics Journal, 28,* 437–453.

Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology, 47,* 1231–1243.

Lord, C., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37,* 2098–2109.

Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs, 54,* 1–95.

Luchins, A. S. (1957). Experimental attempts to minimize the impact of first impressions. In C. I. Hovland (Ed.), *The order of presentation in persuasion.* New Haven, CT: Yale University Press.

Lusted, L. B. (1977). *A study of the efficacy of diagnostic reaiologic procedures: Final report on diagnostic efficacy.* Chicago: Efficacy Study Committee of the American College of Radiology.

Lycan, W. C. (1988). *Judgment and justification.* New York: Cambridge University Press.

MacIntyre, A. (1988). *Whose justice? Which rationality?* Notre Dame, IN: University of Notre Dame Press.

Mackay, C. (1932). *Extraordinary popular delusions and the madness of crowds* (2nd ed.). Boston: Page. (Original second edition published 1852)

Mahoney, M. J. (1976). *Scientist as subject: The psychological imperative.* Cambridge, MA: Ballinger.

Mahoney, M. J. (1977). Publication prejudices. *Cognitive Therapy and Research, 1,* 161–175.

Manktelow, K. I., & Over, D. E. (1990). Deontic thought and the selection task. In K. J. Gilhooly, M. T. G. Keane, R. H. Logic, & G. Erdos. (Eds.), *Lines of thinking* (Vol. 1). London: Wiley.

Manktelow, K. I., & Over, D. E. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition, 39,* 85–105.

Manktelow, K. I., & Over, D. E. (1992). Utility and deontic reasoning: Some comments on Johnson-Laird and Byrne. *Cognition, 43,* 183–188.

Markovits, H., & Savary, F. (1992). Pragmatic schemas and the selection task. *Quarterly Journal of Experimental Psychology, 45A,* 133–148.

Matlin, M. W., & Stang, D. J. (1978). *The Pollyanna principle: Selectivity in language, memory and thought.* Cambridge, MA: Shenkman.

McGuire, W. J. (1960). A syllogistic analysis of cognitive relationships. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm (Eds.), *Attitude organization and change* (pp. 65–110). New Haven, CT: Yale University Press.

Meehl, P. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.* Minneapolis: University of Minnesota Press.

Meehl, P. (1960). The cognitive activity of the clinician. *American Psychologist, 15,* 19–27.

Meichenbaum, D. H., Bowers, K. S., & Ross, R. R. (1969). A behavioral analysis of teacher expectancy effect. *Journal of Personality and Social Psychology, 13,* 306–316.

Merton, R. K. (1948). The self-fulfilling prophecy. *Antioch Review, 8,* 193–210.

Merton, R. K. (1957). Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review, 22,* 635–659.

Millward, R. B., & Spoehr, K. T. (1973). The direct measurement of hypothesis-testing strategies. *Cognitive Psychology, 4,* 1–38.

Mischel, W. (1968). *Personality and assessment.* New York: Wiley.

Mischel, W., Ebbesen, E., & Zeiss, A. (1973). Selective attention to the self: Situational and dispositional determinants. *Journal of Personality and Social Psychology, 27,* 129–142.

Mischel, W., & Peake, P. K. (1982). Beyond *deja vu* in the search for cross-situational consistency. *Psychological Review, 89,* 730–755.

Mitroff, I. (1974). *The subjective side of science.* Amsterdam: Elsevier.

Murphy, A. H., & Winkler, R. L. (1974). Subjective probability forecasting experiments in meteorology: Some preliminary results. *Bulletin of the American Meteorological Society, 55,* 1206–1216.

Murphy, A. H., & Winkler, R. L. (1977). Can weather forecasters formulate reliable probability forecasts of precipitation and temperature? *National Weather Digest, 2,* 2–9.

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

Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1977). Confirmation bias in a simulated research environment: An experimental study of scientific inferences. *Quarterly Journal of Experimental Psychology, 29,* 85–95.

Narveson, R. D. (1980). Development and learning: Complementary or conflicting aims in humanities education? In R. G. Fuller, R. F. Bergstrom, E. T. Carpenter, H. J. Corzine, J. A. McShance, D. W.

Miller, D. S. Moshman, R. D. Narveson, J. L. Petr, M. C. Thornton, & V. G. Williams (Eds.), *Piagetian programs in higher education* (pp. 79–88). Lincoln, NE: ADAPT Program.

Nickerson, R. S. (1996). Hempel's paradox and Wason's selection task: Logical and psychological puzzles of confirmation. *Thinking and reasoning, 2,* 1–31.

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

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

Oskamp, S. (1965). Overconfidence in case study judgments. *Journal of Consulting Psychology, 29,* 261–265.

Over, D. E., & Evans, J. St. B. T. (1994). Hits and misses: Kirby on the selection task. *Cognition, 52,* 235–243.

Pennebaker, J. W., & Skelton, J. A. (1978). Psychological parameters of physical symptoms. *Personality and Social Psychology Bulletin, 4,* 524–530.

Pennington, N., & Hastie, R. (1993). The story model for juror decision making. In R. Hastie (Ed.), *Inside the juror: The psychology of juror decision making* (pp. 192–221). New York: Cambridge University Press.

Perkins, D. N., Allen, R., & Hafner, J. (1983). Difficulties in everyday reasoning. In W. Maxwell (Ed.), *Thinking: The frontier expands.* Hillsdale, NJ: Erlbaum.

Perkins, D. N., Farady, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 83–106). Hillsdale, NJ: Erlbaum.

Peterson, C. R., & DuCharme, W. M. (1967). A primacy effect in subjective probability revision. *Journal of Experimental Psychology, 73,* 61–65.

Pitz, G. F. (1969). An inertia effect (resistance to change) in the revision of opinion. *Canadian Journal of Psychology, 23,* 24–33.

Pitz, G. F. (1974). Subjective probability distributions for imperfectly know quantities. In L. W. Gregg (Ed.), *Knowledge and cognition.* New York: Lawrence Erlbaum Associates.

Pitz, G. F., Downing, L., & Reinhold, H. (1967). Sequential effects in the revision of subjective probabilities. *Canadian Journal of Psychology, 21,* 381–393.

Politzer, G. (1986). Laws of language use and formal logic. *Journal of Psycholinguistic Research, 15,* 47–92.

Polya, G. (1954a). *Mathematics and plausible reasoning. Volume 1: Induction and analogy in mathematics.* Princeton, NJ: Princeton University Press.

Polya, G. (1954b). *Mathematics and plausible reasoning. Volume 2: Patterns of plausible inference.* Princeton, NJ: Princeton University Press.

Popper, K. (1959). *The logic of scientific discovery.* New York: Basic Books.

Popper, K. (1981). Science, pseudo-science, and falsifiability. In R. D. Tweney, M. E. Doherty, & C. R. Mynatt (Eds.), *Scientific thinking* (pp. 92–99). New York: Columbia University Press. (Original work published in 1962)

Price, D. J. de S. (1963). *Little science, big science.* New York: Columbia University Press.

Pyszczynski, T., & Greenberg, J. (1987). Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis-testing model. *Advances in experimental social psychology* (pp. 297–340). New York: Academic Press.

Ray, J. J. (1983). Reviving the problem of acquiescence response bias. *Journal of Social Psychology, 121,* 81–96.

Revlis, R. (1975a). Syllogistic reasoning: Logical decisions from a complex data base. In R. J. Falmagne (Ed.), *Reasoning: Representation and process.* New York: Wiley.

Revlis, R. (1975b). Two models of syllogistic inference: Feature selection and conversion. *Journal of Verbal Learning and Verbal Behavior, 14,* 180–195.

Rhine, R. J., & Severance, L. J. (1970). Ego-involvement, discrepancy, source credibility, and attitude change. *Journal of Personality and Social Psychology, 16,* 175–190.

Rist, R. C. (1970). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. *Harvard Educational Review, 40,* 411–451.

Rosenhan, D. L. (1973). On being sane in insane places. *Science, 179,* 250–258.

Rosenthal, R. (1974). *On the social psychology of self-fulfilling prophecy: Further evidence for Pygmalion effects and their mediating mechanisms.* New York: MSS Modular Publications, Module 53.

Rosenthal, R., & Jacobson, L. (1968) *Pygmalion in the classroom.* New York: Hold, Rinehart and Winston.

Ross, L. (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), *Advances in experimental social psychology, 10,* New York: Academic Press.

Ross, L., & Anderson, C. (1982). Shortcomings in the attribution process: On the origins and maintenance of erroneous social assessments. In A. Tversky, D. Kahneman, & P. Slovic (Eds.), *Judgement under uncertainty: Heursitics and biases.* Cambridge, England: Cambridge University Press.

Ross, L., & Lepper, M. R. (1980). The perseverance

of beliefs: Empirical and normative considerations. In R. Shweder & D. Fiske (Eds.), *New directions for methodology of social and behavioral science: Fallibel judgement in behavioral research* (Vol. 4, pp. 17–36). San Francisco: Jossey-Bass.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perserverance in self perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32,* 880–892.

Ross, L., Lepper, M. R., Strack, F., & Steinmetz, J. L. (1977). Social explanation and social expection: The effects of real and hypothetical explanations upon subjective likelihood. *Journal of Personality and Social Psychology, 35,* 817–829.

Roszak, T. (1986). *The cult of information: The folklore of computers and the true art of thinking.* New York: Pantheon Books.

Salmon, W. C. (1973). Confirmation. *Scientific American, 228*(5), 75–83.

Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological Bulletin, 66,* 178–200.

Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys.* New York: Academic Press.

Schwartz, B. (1982). Reinforcement-induced behavioral stereotypy: How not to teach people to discover rules. *Journal of Experimental Psychology: General, 111,* 23–59.

Sears, D. O., & Freedman, J. L. (1967). Selective exposure to information: A critical review. *Public Opinion Quarterly, 31,* 194–213.

Shaklee, H., & Fischhoff, B. (1982). Strategies of information search in causal analysis. *Memory and Cognition, 10,* 520–530.

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.

Simon, H. A. (1957). *Models of man: Social and rational.* New York: Wiley.

Simon, H. A. (1990). Alternative visions of rationality. In P. K. Moser (Ed.), *Rationality in action: Contemporary approaches* (pp. 189–204). New York: Cambridge University Press. (Original work published 1983)

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.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology, 28,* 1–39.

Smyth, C. P. (1890). *Our inheritance in the great*

pyramid (5th ed.). New York: Randolf. (Original work published 1864)

Snyder, M. (1981). Seek and ye shall find: Testing hypotheses about other people. In E. T. Higgins, C. P. Heiman, & M. P. Zanna (Eds.), *Social cognition: The Ontario symposium on personality and social psychology* (pp. 277–303). Hillsdale, NJ: Erlbaum.

Snyder, M. (1984). When belief creates reality. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18). New York: Academic Press.

Snyder, M., & Campbell, B. H. (1980). Testing hypotheses about other people: The role of the hypothesis. *Personality and Social Psychology Bulletin, 6,* 421–426.

Snyder, M., & Gangestad, S. (1981). Hypotheses-testing processes. *New directions in attribution research* (Vol. 3). Hillsdale, NJ: Erlbaum.

Snyder, M., & Swann, W. B. (1978a). Behavioral confirmation in social interaction: From social perception to social reality. *Journal of Experimental Social Psychology, 14,* 148–162.

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

Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology, 35,* 656–666.

Snyder, M., & Uranowitz, S. W. (1978). Reconstructing the past: Some cognitive consequences of person perception. *Journal of Personality and Social Psychology, 36,* 941–950.

Solomon, M. (1992). Scientific rationality and human reasoning. *Philosophy of Science, 59,* 439–455.

Strohmer, D. C., & Newman, L. J. (1983). Counselor hypothesis-testing strategies. *Journal of Counseling Psychology, 30,* 557–565.

Swann, W. B., Jr., Giuliano, T., & Wegner, D. M. (1982). Where leading questions can lead: The power of conjecture in social interaction. *Journal of Personality and Social Psychology, 42,* 1025–1035.

Taplin, J. E. (1975). Evaluation of hypotheses in concept identification. *Memory and Cognition, 3,* 85–96.

Taylor, J. (1859). *The great pyramid: Why was it built? And who built it?* London: Longman, Green, Longman & Roberts.

Tetlock, P. E., & Kim, J. I. (1987). Accountability and judgment processes in a personality prediction task. *Journal of Personality and Social Psychology, 52,* 700–709.

Thomas, L. (1979). *The Medusa and the snail: More notes of a biology watcher.* New York: Viking Press.

Thurstone, L. L. (1924). *The nature of intelligence.* London: Routledge & Kegan Paul.

Tishman, S., Jay, E., & Perkins, D. N. (1993). Teaching thinking dispositions: From transmission to enculturation. *Theory into practice, 32,* 147–153.

Trabasso, T., Rollins, H., & Shaughnessey, E. (1971). Storage and verification stages in processing concepts. *Cognitive Psychology, 2,* 239–289.

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

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

Trope, Y., Bassock, 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.

Troutman, C. M., & Shanteau, J. (1977). Inferences based on nondiagnostic information. *Organizational Behavior and Human Performance, 19,* 43–55.

Tuchman, B. W. (1984). *The march of folly: From Troy to Vietnam.* New York: Ballantine Books.

Tweney, R. D. (1984). Cognitive psychology and the history of science: A new look at Michael Faraday. In H. Rappart, W. van Hoorn, & S. Bem. (Eds.), *Studies in the history of psychology and the social sciences* (pp. 235–246). Hague, The Netherlands: Mouton.

Tweney, R. D., & Doherty, M. E. (1983). Rationality and the psychology of inference. *Synthese, 57,* 139–161.

Tweney, R. D., Doherty, M. E., Worner, W. J., Pliske, D. B., Mynatt, C. R., Gross, K. A., & Arkkelin, D. L. (1980). Strategies of rule discovery in an inference task. *Quarterly Journal of Experimental Psychology, 32,* 109–123.

Valentine, E. R. (1985). The effect of instructions on performance in the Wason selection task. *Current Psychological Research and Reviews, 4,* 214–223.

Von Däniken, E. (1969). *Chariots of the gods?* New York: Putnam.

Wagenaar, W. A., & Keren, G. B. (1986). Does the expert know? The reliability of predictions and confidence ratings of experts. In E. Hollnagel, G. Maneine, & D. Woods (Eds), *Intelligent decision support in process environments* (pp. 87–107). Berlin: Springer.

Waismann, F. (1952). Verifiability. In A. G. N. Flew (Ed.), *Logic and language.* Oxford: Basil Blackwell.

Wallsten, T. S., & González-Vallejo, C. (1994). Statement verification: A stochastic model of

judgment and response. *Psychological Review,* *101,* 490–504.

Wason, P. C. (1959). The processing of positive and negative information. *Quarterly Journal of Experimental Psychology, 11,* 92–107.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12,* 129–140.

Wason, P. C. (1961). Response to affirmative and negative binary statements. *British Journal of Psychology, 52,* 133–142.

Wason, P. C. (1962). Reply to Wetherick. *Quarterly Journal of Experimental Psychology, 14,* 250.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology I,* Harmondsworth, Middlesex, England: Penguin.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20,* 273–281.

Wason, P. C. (1977). "On the failure to eliminate hypotheses. . . ."—A second look. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking: Readings in cognitive science* (pp. 307–314). Cambridge, England: Cambridge University Press. (Original work published 1968)

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

Wason, P. C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology, 23,* 63–71.

Webster, E. C. (1964). *Decision making in the employment interview.* Montreal, Canada: Industrial Relations Center, McGill University.

Wegener, A. (1966). *The origin of continents and oceans* (4th ed.; J. Biram, Trans.) London: Dover. (Original work published in 1915)

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 39,* 806–820.

Weinstein, N. D. (1989). Optimistic biases about personal risks. *Science, 246,* 1232–1233.

Wetherick, N. E. (1962). Eliminative and enumerative behaviour in a conceptual task. *Quarterly Journal of Experimental Psychology, 14,* 246–249.

Wilkins, W. (1977). Self-fulling prophecy: Is there a phenomenon to explain? *Psychological Bulletin, 84,* 55–56.

Winkler, R. L., & Murphy, A. H. (1968). "Good" probability assessors. *Journal of Applied Meteorology, 1,* 751–758.

Woods, S., Matterson, J., & Silverman, J. (1966). Medical students' disease: Hypocondriasis in medical education. *Journal of Medical Education, 41,* 785–790.

Yachanin, S. A., & Tweney, R. D. (1982). The effect of thematic content on cognitive strategies in the four-card selection task. *Bulletin of the Psychonomic Society, 19,* 87–90.

Zanna, M. P., Sheras, P., Cooper, J., & Shaw, C. (1975). Pygmalion and Galatea: The interactive effect of teacher and student expectancies. *Journal of Experimental Social Psychology, 11,* 279–287.

Ziman, J. (1978). *Reliable knowledge.* Cambridge, England: Cambridge University Press.

Zuckerman, M., Knee, R., Hodgins, H. S., & Miyake, K. (1995). Hypothesis confirmation: The joint effect of positive test stratecy and acquiescence response set. *Journal of Personality and Social Psychology, 68,* 52–60.