

Considering the Opposite: A Corrective Strategy for Social Judgment

Charles G. Lord
Princeton University

Mark R. Lepper
Stanford University

Elizabeth Preston
Princeton University

It is proposed that several biases in social judgment result from a failure—first noted by Francis Bacon—to consider possibilities at odds with beliefs and perceptions of the moment. Individuals who are induced to *consider the opposite*, therefore, should display less bias in social judgment. In two separate but conceptually parallel experiments, this reasoning was applied to two domains—biased assimilation of new evidence on social issues and biased hypothesis testing of personality impressions. Subjects were induced to consider the opposite in two ways: through explicit instructions to do so and through stimulus materials that made opposite possibilities more salient. In both experiments the induction of a consider-the-opposite strategy had greater *corrective effect* than more demand-laden alternative instructions to be as fair and unbiased as possible. The results are viewed as consistent with previous research on perseverance, hindsight, and logical problem solving, and are thought to suggest an effective method of retraining social judgment.

“I beseech ye in the bowels of Christ, think that ye may be mistaken.” I should like to have that written over the portals of every church, every school, and every courthouse, and, may I say, of every legislative body in the United States.” Thus spoke Judge Learned Hand in 1951, so taken was he with the wisdom of Oliver Cromwell’s 1650 plea to the Church of Scotland. The criticism that human decision makers do not adequately consider alternative possibilities, especially those directly at odds with their beliefs and perceptions of the moment, remains as viable today as it was in Cromwell’s time. In fact, modern psychology has provided substantial empirical evidence to buttress the argument that our beliefs pervasively color and bias our response to subsequent information, evidence, or argumentation (e.g., Allport, 1954; Asch, 1946; Kahneman, Slovic, & Tversky, 1982;

Nisbett & Ross, 1980; Ross & Lepper, 1980; Snyder, 1981).

Cromwell’s plea is, of course, a very general admonition that could be interpreted as an exhortation to try harder—a caution that would imply a motivational account of human fallibility and a largely motivational prescription for more rational judgment. Raise the stakes, as the United States did in Vietnam, and the other side will begin to view the issue more rationally (Tuchman, 1984). The success of such appeals in history and in current research, however, suggests that merely trying harder may be less than a foolproof debiasing strategy (cf. Kahneman et al., 1982; Nisbett & Ross, 1980).

We believe that there are also more specific and more cognitive elements involved in this characteristic failure to *consider the opposite* and that these processes may underlie many attributional and judgmental errors. In particular, we would argue, people typically seem oblivious to the fact that the way they process information may itself influence their judgments and that the questions they ask may determine the answers they receive. Thus any inducement for decision makers to consider that matters might be other than what they seem, especially an inducement to consider

This research was supported in part by National Institute of Mental Health Grant MH-36093 to Mark R. Lepper and Lee Ross. We thank Lee Ross for comments on earlier drafts and Mark Snyder for making available stimulus materials for Experiment 2.

Requests for reprints should be sent to Charles Lord, Department of Psychology, Princeton University, Princeton, New Jersey 08540.

possibilities diametrically opposed to one's assumptions, would have an ameliorative effect on judgmental bias. Judge Hand's suggestion, in short, might be taken seriously by those interested in promoting more rational social judgment.

Such a strategy might be implemented in several ways. One general approach might involve direct instructions to consider various hypothetical and opposite possibilities; for example, when a professor asks a new graduate student to consider what the data from a proposed experiment might mean if the expected results were reversed. This approach is *direct*, in that the professor describes the tendency to overlook alternative data patterns and explicitly instructs the student to imagine these outcomes. A second general approach might be to alter the task or eliciting stimulus conditions in such a way as to make opposite possibilities more salient; for example, when the professor merely asks the student to read a paper whose conclusions suggest an experimental outcome opposite to that expected by the student who has read only one side of a theoretical dispute. This approach is *indirect*, in that the professor neither describes the tendency to ignore alternative data patterns nor instructs the student to adopt any particular cognitive strategy, but instead relies on the recommended paper to render opposite possibilities more accessible. In the present studies, we sought to induce consideration of opposite possibilities in two ways: directly, through explicit instructions, and indirectly, through stimulus salience and increased accessibility. Both the direct and the indirect approaches were compared with an alternative manipulation that reflected the different assumption that biased judges are insufficiently motivated.

In order to test the generality of considering the opposite as a debiasing strategy, we applied it to two different domains of social judgment: biased assimilation of new evidence (Lord, Ross, & Lepper, 1979) and biased hypothesis testing (Snyder & Swann, 1978). We chose these two domains deliberately because they seemed more involving than many statistical or mathematical problems such as probability or covariation estimation (Jennings, Amabile, & Ross, 1982; Kahneman & Tversky, 1972; 1973) and thus presumably more resistant to correction.

Experiment 1

Biased assimilation of new evidence serves as a good example of what can happen when opposite possibilities are overlooked. Those who hold strong beliefs about an issue are apt to examine relevant evidence in a biased manner, by accepting confirming evidence at face value and subjecting disconfirming evidence to highly critical evaluation (Lord et al., 1979). As a result, partisans on both sides of an issue may adopt more extreme attitudes following exposure to mixed evidence, some of it supporting one side of an issue and some the other.

Lord et al. (1979) asked subjects who either supported or opposed capital punishment to read two purported studies, one seemingly confirming and one seemingly disconfirming the subject's beliefs about the deterrent efficacy of the death penalty. Both proponents and opponents of capital punishment rated those procedures that produced confirming results as methodologically superior to those that produced disconfirming results, and both used this perceived disparity in the quality of evidence on the two sides of the issue as justification for adopting more polarized attitudes. The researchers concluded that attempts to furnish objective evidence on burning social issues "will frequently fuel rather than calm the fires of debate" (1979, p. 2108).

For those who value social science evidence on complex and important social issues, the way in which Lord et al.'s (1979) subjects evaluated new evidence seems less than optimal. We ought, therefore, to be interested in ways to inhibit an uncritical biased assimilation of new evidence to existing beliefs and attitudes. The appropriate method of correction, however, depends on where one believes the bias to lie. One possibility is that the subjects in Lord et al.'s (1979) study were not sufficiently motivated to be honest, accurate, and unbiased and were not prepared to suspend judgment until they could give equal consideration to both sides, as jurors in the legal setting and elected representatives in the legislative setting are often reminded to do. The remedy suggested by this analysis is to instruct and educate prospective decision makers in the exercise of impartiality. A second possibility is suggested by our earlier

analysis. Thus Lord et al.'s (1979) subjects may have responded to a study's methodology on the basis of its stated result, without considering the possibility that the same methodology might have produced an opposite conclusion. The remedy suggested by this analysis is to promote an explicit consideration of alternative possibilities, especially those possible outcomes that are diametrically opposed to those expected or perceived. Experiment 1 tested both the "be unbiased" and the "consider-the-opposite" remedies in a replication of Lord et al.'s (1979) study on biased assimilation of new evidence.

Method

One hundred twenty Stanford University undergraduates participated in partial fulfillment of a course requirement. Twenty proponents and twenty opponents of capital punishment received each of three types of instructions.

In a replication condition we used the subject selection criteria, experimental materials, and procedure described in greater detail by Lord et al. (1979). We selected as subjects students who on an earlier questionnaire had either favored capital punishment and believed that it deterred potential murderers (proponents) or opposed capital punishment and believed that it did not deter potential murderers (opponents). In a 1-hr laboratory session, each student received four pieces of information: first, a one-sentence summary of a purported empirical result demonstrating the death penalty's effectiveness or ineffectiveness in lowering murder rates; second, a two-page description of the methodology that produced this result; third, a one-sentence summary of an empirical result opposite to that found in the first study; fourth, a two-page description of the methodology that produced this second result. After reading each of the four pieces of information, subjects indicated how much and in what direction their attitudes toward capital punishment and their beliefs about its deterrent efficacy had changed, both as a result of that piece of information alone and cumulatively. In addition, after reading each of the two-page descriptions, subjects rated how well done (from $-8 = \text{very poorly done}$ to $8 = \text{very well done}$) and how convincing (from $-8 = \text{completely unconvincing}$ to $8 = \text{completely convincing}$) the described study seemed as evidence on the issue. The overall design was counter-balanced with respect to subjects' initial attitudes, order of confirming versus disconfirming information, and which methodology was said to have produced which result.

In a be-unbiased condition we added to the replication instructions a warning that "the particular studies you select¹ may provide evidence on the same side of this issue in both cases, or they may provide evidence on different sides of the issue," and continued:

We would like you to be as *objective* and *unbiased* as possible in evaluating the studies you read. You might consider yourself to be in the same role as a judge or juror asked to weigh all of the evidence in a fair and impartial manner."

In a consider-the-opposite condition we described the process by which biased assimilation is thought to occur (e.g., that strengths and weaknesses may be differentially salient), and recommended the following:

Ask yourself at each step whether you would have made the same high or low evaluations had exactly the same study produced results on the *other* side of the issue.

One way of characterizing the difference between the be-unbiased and consider-the-opposite instructions is that subjects in the former condition were told, "Here's what can happen. Don't let it happen to you," whereas subjects in the latter condition were told, "Here's how it happens and what you can do about it." Consider-the-opposite instructions were thus analogous to Ross, Lepper, and Hubbard's (1975) successful technique of overcoming perseverance by describing how it happens and reminding subjects that a different experimental experience might have brought different supporting cognitions to mind. Merely describing a bias, at least in an involving domain, has no ameliorative effect (Fischhoff, 1977, 1982), so the operative component of consider-the-opposite instructions was assumed to be the recommended strategy.

Results and Discussion

Evaluations. Lord et al. (1979) found preferential evaluations of how well done and how convincing the confirming and disconfirming studies seemed and subsequent attitude polarization. We examined the same measures in order to test whether the three different types of instructions had different effects. More specifically, we conducted a 3×2 (Condition: Replication, Be-Unbiased, Consider-the-Opposite \times Initial Attitude: Proponent, Opponent) analysis of variance (ANOVA) of differences between subjects' evaluations of the antideterrence and prodeterrence studies. The results are presented in Table 1.

As shown by the pattern of difference scores in Table 1, instructions interacted with initial attitude in determining evaluations of how well done the studies were, $F(2, 114) = 4.21, p < .05$. Initial attitude made a difference for students who received the replication instructions, $F(1, 114) = 6.65, p < .05$, proponents finding the prodeterrence study better done than the antideterrence study ($M = 1.6$) and opponents finding the prodeterrence study worse done ($M = -1.1$). Initial attitude also made a difference for students who received

¹ As described in Lord et al. (1979, p. 2100), subjects "chose" the two studies that they were to read from a set of 10 that were in reality identical.

Table 1
Mean Evaluations of Prodeterrence and Antideterrence Studies by Proponents and Opponents of Capital Punishment as a Function of Instructions in Experiment 1

Instructions	Study	How well done?		How convincing?	
		Proponents	Opponents	Proponents	Opponents
Replication	Prodeterrence	.8	-.6	1.5	-.8
	Antideterrence	-.8	.5	-1.4	.2
	Difference	1.6	-1.1	2.9	-1.0
Be-unbiased	Prodeterrence	1.7	-1.6	1.6	-2.5
	Antideterrence	-.7	.1	-1.6	1.0
	Difference	2.4	-1.7	3.2	-3.5
Consider-the-opposite	Prodeterrence	-.3	.4	.8	.2
	Antideterrence	-.6	-.1	-.2	.4
	Difference	.3	.5	1.0	-.2

Note. Positive difference scores indicate prodeterrence study better done/more convincing; negative difference scores indicate antideterrence study better done/more convincing.

the be-unbiased instructions, $F(1, 114) = 15.51$, $p < .01$, proponents finding the prodeterrence study better done ($M = 2.4$) and opponents finding the prodeterrence study worse done ($M = -1.7$). Initial attitude, however, did not affect the evaluations of students who received consider-the-opposite instructions, $F(1, 114) < 1$. To compare the effects of the three types of instructions directly, we conducted the same 3×2 analysis for difference scores that reflected preference for attitude-confirming evidence (pro minus anti for proponents; anti minus pro for opponents). According to a Newman-Keuls test following this analysis, consider-the-opposite instructions produced significantly less attitude-congruent evaluations than either replication or be-unbiased instructions, which did not differ ($p < .05$).

As also shown in Table 1, instructions interacted with initial attitude in determining evaluations of how convincing the studies seemed as evidence on the issue of capital punishment, $F(2, 114) = 3.95$, $p < .05$. On this measure as well, initial attitude made a difference for students who received the replication instructions, $F(1, 114) = 8.13$, $p < .01$, proponents finding the prodeterrence study more convincing than the antideterrence study ($M = 2.9$) and opponents finding the prodeterrence study less convincing ($M = -1.0$). Initial attitude also made a difference for students who received the be-unbiased instructions, $F(1, 114) = 23.76$, $p < .01$, proponents finding the prodeterrence study more

convincing ($M = 3.2$) and opponents finding the prodeterrence study less convincing ($M = -3.5$). Initial attitude, however, did not affect the evaluations of students who received consider-the-opposite instructions, $F(1, 114) < 1$. As with the well-done measure, we conducted the same 3×2 analysis for difference scores that reflected a tendency to find attitude-congruent evidence more convincing than attitude-incongruent evidence. According to a Newman-Keuls test following this analysis, consider-the-opposite instructions produced less attitude-congruent evaluations than be-unbiased instructions, with neither differing significantly from replication instructions, ($p < .05$).

Attitude polarization. The striking consequence of subjects' differential evaluations of confirmatory versus disconfirmatory research, Lord et al. (1979) demonstrated, was increased polarization of partisans' attitudes toward capital punishment following exposure to both positive and negative results. Thus, we also examined reported attitude changes from the experiment's start to its finish in subjects' beliefs about the death penalty's deterrent efficacy and in their attitudes on capital punishment. The primary question was whether instructions to consider the opposite would produce not only less biased evaluations of the relevant evidence but also less subsequent belief and attitude polarization. Figure 1 displays the results graphically as mean deviations from a central line that represents no attitude change. The graph

collapses across subjects who read a pro-deterrence study first and an antideterrence study second and those who read the same studies in the other order, and depicts only attitude change following the second (and last) study, regardless of which it was. Positive changes indicate that the net result of reading both studies was a shift toward greater belief in the death penalty's deterrent efficacy or a

more positive attitude toward capital punishment; negative changes indicate that the net result of reading both studies was a shift toward less belief in the death penalty's deterrent efficacy or a more negative attitude toward capital punishment.

As may be seen in the top panel of the figure, after reading the summary and description of both studies, subjects in the

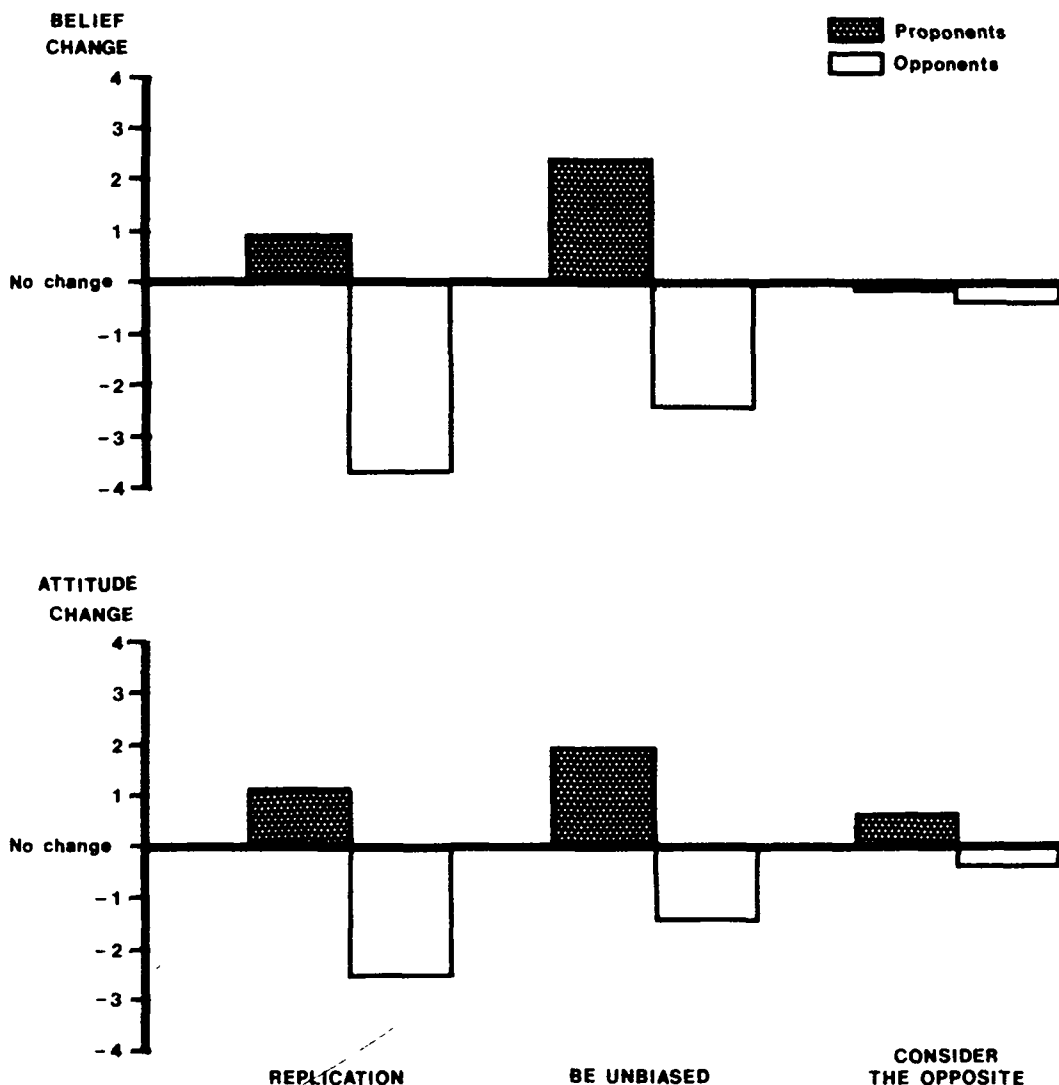


Figure 1. Mean changes in beliefs and attitudes in response to prodeterrence and antideterrence studies by proponents and opponents as a function of instructions in Experiment 1. Belief change ordinate reflects increased (positive numbers) or decreased (negative numbers) belief that the death penalty deters potential murderers. Attitude change ordinate similarly reflects more or less favorable attitude toward capital punishment.

replication condition reported that they had become more extreme in their beliefs about deterrent efficacy [proponents .9, opponents -3.7; $t(39) = 4.93$, $p < .001$], as did subjects admonished to be unbiased [proponents 2.3, opponents -2.4; $t(39) = 4.04$, $p < .001$], but subjects asked to consider the opposite did not [proponents -.1, opponents -.4; $t(39) < 1$]. This pattern of responses produced a significant Instructions \times Initial Attitude interaction, $F(2, 114) = 8.39$, $p < .01$. According to a Newman-Keuls test, consider-the-opposite instructions produced significantly less belief polarization on the deterrent efficacy question than either replication or be-unbiased instructions, which did not differ ($p < .05$).

Similar results characterized reported attitude change. As shown in the bottom panel of Figure 1, after reading the summary and description of both studies, subjects in the replication condition reported that they had shifted to an attitude more extreme than their initial attitude [proponents 1.1, opponents -2.5; $t(39) = 5.35$, $p < .001$], as did subjects admonished to be unbiased [proponents 1.9, opponents -1.4; $t(39) = 3.29$, $p < .01$], but subjects asked to consider the opposite displayed no such attitude polarization [proponents .6, opponents -.4; $t(39) < 1$]. This pattern of responses yielded a significant Instructions \times Initial Attitude interaction, $F(2, 114) = 6.57$, $p < .01$. Again, the three types of instructions differed significantly in their effects on attitude polarization, in that consider-the-opposite instructions produced less attitude polarization than either replication or be-unbiased instructions, which did not differ, according to a Newman-Keuls test ($p < .05$).

Demand characteristics. We believed that the unbiased evaluations and lack of attitude polarization found with consider-the-opposite instructions were a direct result of providing subjects with a *corrective strategy* for social judgment. Another possibility, however, was that the description of biased assimilation and the injunction to "ask yourself, at each step, whether you would have made the same high or low evaluations had exactly the same study produced results on the *other* side of the issue" were laden with demand characteristics. Perhaps subjects felt that these in-

structions put pressure on them to claim lack of bias and attitude change, even though they actually viewed the evidence against them as weak compared to the evidence supporting their own initial attitudes.

To test this possibility, we showed 20 different Stanford undergraduates photocopies of the three types of instructions and asked them to rate the three conditions of Experiment 1 on "how much each type of instructions made it seem that we, the experimenters, would *like* the subjects to report that the two studies, one supporting their initial attitude and the other contradicting their initial attitude, were equally well done," on a scale from 0 = *absolutely no pressure* to 3 = *a lot of pressure*. The students used similar scales to rate each type of instruction on pressure to report unchanged attitudes. For each of these ratings they were also asked to choose "the one set of instructions that applied the *most* pressure of this sort."

These undergraduate raters viewed the replication instructions as embodying the least demand to claim that the studies had been equally well done ($M = .2$), the consider-the-opposite instructions a medium amount ($M = 1.8$), and the be-unbiased instructions the most ($M = 2.7$), $F(2, 38) = 121.83$, $p < .001$. The three means all differed from each other by Newman-Keuls test ($p < .05$). In addition, 18 of the 20 raters indicated that the be-unbiased instructions contained the most demand characteristics of the three, $\chi^2 = 29.21$, $p < .01$. Approximately the same pattern was found in ratings of pressure to claim unchanged attitudes. The raters viewed the replication instructions as embodying the least demand of this sort ($M = .6$), and the consider-the-opposite and be-unbiased instructions as entailing considerably more pressure ($M = 1.60$ and 1.65 , respectively), $F(2, 38) = 6.46$, $p < .01$.

Thus, the corrective effect of consider-the-opposite instructions may not be attributed simply to demand characteristics, or else the be-unbiased instructions would have been at least as effective in seeming to overcome biased assimilation and attitude polarization. Our conclusion was that biases in social judgment can be corrected only by a change in strategy, not just by investing greater effort in a strategy that led to biased judgments in

the first place. Debriefing conversations in the present experiment and in Lord et al.'s (1979) suggested that subjects who were not provided with an alternative judgmental strategy believed that they were being accurate and unbiased. Exhorting them to do more of the same should have had no corrective effect, and it did not.

Experiment 2

The consider-the-opposite technique proved effective, in Experiment 1, in eliminating biased assimilation, but this strategy's benefits may well have been limited to just this one type of judgmental bias. A more convincing case for the strategy's generality would depend on its success in overcoming a different bias of social judgment. In addition, one would want to prompt consideration of opposite possibilities without using direct experimenter instructions to do so. From several possible candidate domains, we selected biased hypothesis testing—the tendency to seek more avidly evidence that promises to confirm than evidence that promises to disconfirm one's hypotheses. Whereas biased assimilation, as investigated in Experiment 1, involves a preferential treatment of new information presented to the individual, biased hypothesis testing involves a more active preferential search for new information.

Snyder and Swann (1978) provided evidence that biased hypothesis testing affects impression formation. They told student subjects that they would interact with another student. Some subjects were asked to test the hypothesis that the other student was an extravert; other subjects were asked to test the hypothesis that the other student was an introvert. Both planned to test their hypotheses by preferentially eliciting confirming information. To test the extravert hypothesis, they wanted to ask the other student questions like "What would you do if you wanted to liven things up at a party?" To test the introvert hypothesis, they wanted to ask the other student questions like "What factors make it hard for you to really open up to people?" These are obviously leading questions that could make almost any respondent seem to confirm the hypothesis.

Subsequently, a carefully conducted series of experiments has shown that biased hy-

pothesis testing in impression formation is an extremely difficult tendency to overcome. Snyder (1981) attempted to undo the bias by informing subjects that the hypothesis was merely hypothetical, by making the hypothesis seem implausible, by rewarding accuracy, by rephrasing affirmative descriptions in the negative, and so on—all to no avail.

We suspected that biased hypothesis testing can result from a blind spot about opposite possibilities.² Specifically, subjects in Snyder and Swann's (1978) studies may have considered only the possibility that the other student had characteristics associated with the hypothesized trait, characteristics that could be tapped by the confirmatory questions they posed. They may never have stopped to consider the possibility that asking the opposite questions might have confirmed an opposite hypothesis, because the opposite characteristics did not spontaneously come to mind. We therefore sought to replicate as nearly as possible the original study, but with a change in experimental context that we hoped would induce subjects on their own, without explicit instructions from the experimenter, to consider the possibility that the other student's personality might prove exactly opposite to their expectations. We suspected that, just as in Experiment 1, exhortations to be more fair, accurate, and unbiased would have no effect on the outcome. Indeed, Snyder (1981) had tried a variety of motivators, including large cash prizes for the most diagnostic questions (Snyder & Swann, 1978, Experiment 4), none of which had a corrective effect. To be certain, however, we included a be-unbiased condition in order to enhance comparison with our first experiment.

Method

Thirty Princeton undergraduates were paid for their participation. Ten received each of three types of instructions.

² Trope and Bassock (1982) suggested that subjects left to their own devices would not frequently employ leading questions of the type used in Snyder and Swann's (1978) experiments but would prefer more diagnostic questions if available. They also noted that subjects in Snyder and Swann's (1978) experiments might not have realized what characteristics would be opposite to those that they had recently read. The latter contention is similar to our belief that subjects failed to consider the opposite.

In a replication condition we used the experimental materials and procedure described in greater detail by Snyder and Swann (1978, Experiment 1). We told subjects that they were to attempt to find out about another person, who was supposedly waiting in another room, by choosing 12 questions from a list of 26 "topic areas often covered by interviewers." Eleven of the questions solicited information about extraverted behaviors (e.g., "In what situations are you most talkative?"), 10 questions solicited information about introverted behaviors (e.g., "In what situations do you wish you were more outgoing?"), and 5 questions were neutral (e.g., "What are your career goals?"). Because Snyder and Swann had found symmetrical biased testing of both introvert and extravert hypotheses, we used only one, the latter. We asked the subjects to choose 12 questions that would be most relevant and informative for deciding whether the other person was an extravert. As an aid in this task, we provided each subject with a "personality profile" of the typical extravert. This was the same profile that Snyder and Swann (1978) had provided for their subjects.

The instructions for the be-unbiased condition were identical to those in the replication condition, except that we added:

Remember that we want you to find out as *accurately* as you can whether the person you are talking to is an extravert or not. The questions you are choosing from have been rated by clinical psychologists on how much insight each provides into a person's character. In other words, we want you to be as *accurate* as possible in providing a *fair* and *unbiased* test of the person's true character.

The instructions for the consider-the-opposite condition were identical to those in the replication condition, except that the experimenter claimed to be unable to locate the extravert profile normally used, and provided instead the introvert profile used by Snyder and Swann (1978), commenting that "Introverts are the opposite of extraverts, so reading this profile should be just as helpful to you." The experimenter later reminded these subjects that although they had read a description of a typical introvert, their task was still to determine whether or not the other student was an extravert. This was an indirect approach in that the experimenter neither described how biased assimilation occurs nor recommended a specific cognitive strategy. We relied instead on stimulus salience to render opposite possibilities more cognitively accessible.

Results and Discussion

Biased hypothesis testing. As shown in Table 2, subjects in the replication condition planned to ask more extravert questions ($M = 7.8$) than introvert questions ($M = 2.2$), $t(9) = 5.92$, $p < .01$. Subjects who were instructed to be unbiased and accurate also planned to ask more extravert questions ($M = 6.7$) than introvert questions ($M = 3.3$), $t(9) = 4.07$, $p < .01$. Subjects induced to consider the opposite by reading an opposite

Table 2

Mean Number of Extravert, Introvert, and Neutral Questions as a Function of Instructions and Procedure in Experiment 2

Instructions	Question type		
	Extravert	Introvert	Neutral
Replication	7.8	2.2	2.0
Be unbiased	6.7	3.3	2.0
Consider the opposite	5.6	4.3	2.0

personality profile showed no such preference for hypothesis-confirming information ($M = 5.6$ and 4.3 , $t < 1$). Thus a 3×3 (Instruction Type \times Question Type) ANOVA of these data yielded a marginally significant interaction, $F(2, 27) = 2.98$, $p < .10$. Most important, subjects asked to consider the opposite were less likely to prefer extravert to introvert questions than were subjects in the replication condition, $t(18) = 2.21$, $p < .05$.

Demand characteristics. As in Experiment 1, in order to rule out explanations based entirely on experimental demand, we described the three types of instructions and procedures to 24 Princeton undergraduates and asked them to rate each condition on the extent to which "the procedure we used made it seem that we *wanted* subjects in that condition to select equal numbers of extravert and introvert questions." On a scale from 0 = *absolutely no pressure* to 3 = *a lot of pressure*, these student raters viewed the replication instructions and procedure as embodying the least demand ($M = .7$), the consider-the-opposite instructions and procedure more ($M = 1.5$), and the be-unbiased instructions and procedure most ($M = 2.2$), $F(2, 36) = 13.00$, $p < .001$. The three means all differed significantly from each other by Newman-Keuls test ($p < .05$). In addition, 19 of 24 raters indicated that the be-unbiased instructions and procedure contained the most demand characteristics of the three, $\chi^2 = 24.25$, $p < .001$. As in attempts to overcome biased assimilation, the corrective effect of consider-the-opposite instructions may not be attributed merely to demand characteristics, or else the be-unbiased instructions and procedure would have been even more effective than the consider-the-opposite instructions in eliminating the pref-

erence for information likely to confirm hypotheses about another person's characteristics.

General Discussion

The results of these two experiments suggest that Judge Hand's assessment of the general utility of Cromwell's admonition was correct. In two different domains of social judgment, biased assimilation of new evidence and biased hypothesis testing, and with two different inducements, direct instructions and indirect manipulation of accessibility through stimulus salience, the cognitive strategy of considering opposite possibilities promoted impartiality.³ In each case, this strategy had more of a corrective effect than more strongly demand-laden exhortations to be fair, accurate, and unbiased. These findings are generally congruent with the proposition that many biases of social judgment are the result of inadequate cognitive strategies rather than inadequate motivation. No matter how hard a person tries to solve a puzzle, the answer will often be found only through breaking set and adopting a new strategy.

The observation that humans have a blind spot for opposite possibilities is not a new one. In 1620, Francis Bacon wrote that "it is the peculiar and perpetual error of human intellect to be more moved and excited by affirmatives than by negatives." Smoke (1933) demonstrated empirically that concepts are more difficult to learn when they are instantiated by negative than by positive instances. Thus, like the inspector in *The Silver Blaze* (Doyle, 1893/1974), we do not consider especially informative the fact that the dog did *not* bark during the night (see Nisbett & Ross, 1980, p. 48). Although it is possible to learn from negative examples and nonoccurrences, they are usually little noticed and unlikely to be taken very seriously even when called to our attention (Hovland & Weiss, 1953; Jenkins & Sainsbury, 1970). This greater reliance on affirmatives than on negatives even affects the attitudinal inferences that we draw from observing our own behavior (Bem, 1972; Fazio, Sherman, & Herr, 1982).

Bacon (1620/1960) also suggested a remedy, an alternative strategy that, if followed,

would help overcome the bias toward positive instances. He wrote admiringly of a man who, when shown temple paintings of individuals who had "paid their vows" to the gods and then survived shipwreck, inquired "But where are they painted that were drowned, after their vows?" Many superstitions develop because the observers of a miracle, for example, an eclipse of the sun that terminates after a virgin is thrown into a volcano, fail to consider and never dare to test what might have happened had the action not been taken. In fact, modern philosophers of science believe that the search for knowledge would proceed more surely were we to abandon our reliance on positive or confirming experimental results and emphasize instead negative or falsifying data (Platt, 1964; Popper, 1965).

Evidence from Other Judgment Domains

In the present experiments, the strategy of considering the opposite eliminated two robust judgmental biases—biases in domains that elicit high personal involvement. The strategy succeeded where admonitions to be fair and unbiased failed, even though it seemed to entail less demand that subjects "mend their ways." We claim that the consider-the-opposite strategy proposed by Bacon and tested here in two different domains of social judgment and with two different methods of instantiation, has been the operative component of successful attempts to overcome bias in several other personally involving judgmental domains, including perseverance, hindsight, and logical problem solving.

Perseverance is the tendency to retain existing beliefs even after the original evidence

³ *Impartiality* is a relative term (Swann, 1984). In some (perhaps less involving) circumstances, considering the opposite could result in overweighting of disconfirmations, a different kind of partiality. Note also that subjects in Experiment 1 were justified in questioning the methodological rigor of a study that produced the "wrong answer," just as one would be justified in questioning the methodology behind a report that women earn more than men for similar work or that eating salt alleviates hypertension. It is only allowing such (often justifiably) negative evaluations to polarize attitudes that is an error, in that it represents circular reasoning. (See Lord et al. 1979, pp. 2106–2107, for a fuller discussion of the normative issue.)

that fostered those beliefs has been shown to be invalid⁴ (Ross et al. 1975). Explanations of perseverance have focused on the otherwise dormant consonant cognitions that the original evidence brought to mind. Subjects who succeeded at a task, for example, would remember previous related successes, whereas subjects who failed would remember previous related failures (Ross et al. 1975). Successful attempts to undo perseverance have required subjects to construct causal explanations for relations opposite to those indicated by the original evidence (Anderson, 1982; Anderson, Lepper, & Ross, 1980; Ross et al. 1975). Thus the exercise of considering opposite possibilities has been shown to reduce unwarranted belief perseverance.

Hindsight is the tendency to exaggerate what could have been anticipated in foresight. As compared to controls, who are not told the correct answer to a question in advance, subjects consistently overestimate in retrospect their likelihood of having been correct; their hindsight is better than their foresight (Fischhoff, 1975). Researchers have tested a wide variety of corrective techniques, the only successful ones of which have involved explicitly asking subjects to write reasons for the wrong answer or for the outcome that did not happen (Koriat, Lichtenstein, & Fischhoff, 1980; Slovic & Fischhoff, 1977). The successful procedure seems analogous to the direct consider-the-opposite approach used in the present Experiment 1.

Logical problem solving involves deciding whether an abstract conditional rule is true or false (Wason, 1966; Wason & Johnson-Laird, 1972). Subjects may, for example, be shown a deck of cards that all have a letter on one side and a number on the other, and asked which cards must be turned over to determine the truth or falsity of a conditional rule such as "if the letter is a vowel then the number on the other side is odd." Subjects usually neglect to turn cards whose exposed side shows an even number. Of the various "remedies" attempted, the most effective has been to ask subjects to consider what might be on the other side of the overlooked correct choice (that is, the card with a "6" might have a vowel on the other side). This consider-the-opposite procedure was even more effective when the experimenter physically turned

a neglected card to reveal that it falsified the rule (Wason, 1969; Wason & Golding, 1974), a technique that seems analogous to that used in the present Experiment 2, in which stimulus salience was used to induce subjects to consider "the other side of the coin" in their search for confirming and disconfirming personality characteristics.

Mechanisms

One important question is why strategies of the consider-the-opposite type should have a corrective effect in judgment domains as involving as biased assimilation of new evidence on controversial issues and as involving as biased hypothesis testing about the personality of another student with whom the subject expects to interact. One explanation involves affect. Subjects in Lord et al.'s (1979) biased assimilation demonstration, for example, might well have had an initial positive affective reaction to a confirming result and an initial negative affective reaction to a disconfirming result. The affect attached to the result might have become associated with the methodology that produced the result. Subjects in Experiment 1 who were asked to consider their reaction "had the same methodology produced an opposite result," may have had this affective connection broken by an antagonistic affective reaction. We do not dismiss this possibility, especially in a domain so affect-laden as attitudes about capital punishment, but it seems a less plausible explanation of Experiment 2, where subjects were less likely to attach a strong affective reaction to either the extravert or the introvert personality profile. Much less does this "affective antagonism" explanation seem able to account for parallel findings in perseverance, hindsight, or logical problem solving.

A second explanation of the strategy's effectiveness involves anchoring. Tversky and Kahneman (1974) have proposed that judgments under uncertainty are often affected by the starting point from which they are reached. Subjects asked to estimate the pop-

⁴ This is different from biased assimilation of new evidence (Lord et al., 1979), in which new evidence is added to a belief rather than old evidence subtracted.

ulation of a city of less than one million will guess a larger number than will those asked to estimate the population of a city of more than one hundred thousand, even though the city involved is the same. Jones (1979) and Quattrone (1982) have suggested that anchoring may lie at the heart of the "fundamental attribution error," the tendency to attribute the behavior of others to dispositional causes even when external constraints are obviously operative. When we see someone behave, we use that behavior as an anchor estimate of their true intentions, and then adjust insufficiently for situational constraints. If anchoring is the major mechanism in the fundamental attribution error, then asking subjects to imagine the actor behaving in an opposite way might prove an effective therapy.

Perhaps the most reasonable explanation of the consider-the-opposite strategy's success involves both anchoring and construct accessibility.⁵ In a series of elegant experiments, Higgins and his colleagues (Higgins & King, 1981; Higgins, Rholes, & Jones, 1977) have developed the idea that social judgments are influenced importantly by the constructs that have recently been activated or primed (see also Taylor and Fiske, 1978, on "top of the head phenomena"). Social perceivers who have recently thought about hostility, for example, are more apt to interpret an individual's ambiguous behavior as hostile (Higgins et al. 1977), and problem solvers who have recently thought about containers as separate from their contents (a carton and eggs versus a carton of eggs) are more apt to think of using a box as a makeshift platform (Higgins & Chaires, 1980). The consider-the-opposite strategy may well make the opposite anchor as accessible as that suggested by immediate experience. The possibility of a negative result in the biased assimilation of new evidence, of opposite personality characteristics in biased hypothesis testing of personality impressions, of a diametrically opposed causal sequence in perseverance, of an alternative answer proving correct in hindsight, and of a falsifying instance in logical problem solving, are all made more accessible by manipulations belonging to what we see as the consider-the-opposite family of judgmental strategies. Although the argument is only speculative at present, to us the mechanism that underlies

the corrective effect of considering the opposite is most likely one of anchor accessibility.

Retraining

We cannot but conclude that Judge Hand's advice should be taken literally. Nisbett and Ross (1980), in their influential review of biases in social judgment, concluded that human judgment could be improved by teaching statistics at earlier grade levels. Indeed, some errors of judgment under uncertainty are less apt to be made by subjects with formal statistical training (Kahneman et al., 1982). It is possible, however, that the ameliorative effects of statistical training depend importantly on internalization of a consider-the-opposite rule for judgments in general. Learning to be wary of and to identify Type I errors, for example, ought to encourage the consideration of opposite possibilities. Perhaps direct training in the logical technique of *modus tollens*, or guided experience in avoiding errors by considering the opposite, might have more effect than statistical training, especially in the realm of social judgments, which are often more involving than mathematical puzzles or statistical problems and thus presumably more resistant to change.

The most effective form of retraining for social judgment, then, may involve a change in strategy rather than in motivation. In many psychological experiments, our subjects may be trying to solve a different problem than the one that we believe we have posed (Henle, 1962), or may not hit on the most effective strategy unless it is made apparent to them either through explicit instructions or through contextual salience. One possible remedy for biases in social judgment is to induce greater use of the consider-the-opposite strategy, a double-check procedure that seems necessary if we are to overcome the blind spot that Bacon (1620/1960) referred to as a "peculiar and perpetual error of human intellect."

⁵ Interestingly, subjects who read both the introvert and extravert profiles still engage in biased hypothesis testing (Snyder & Campbell, 1980), which suggests that the initial anchor has preference unless it is totally supplanted by a new and opposite anchor as in the present Experiment 2.

References

- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Anderson, C. A. (1982). Inoculation and counterexplanation: Debiasing techniques in the perseverance of social theories. *Social Cognition*, 1, 126-139.
- Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology*, 39, 1037-1049.
- Asch, S. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 258-290.
- Bacon, F. (1960). *The new organon and related writings*. New York: Liberal Arts Press. (Original work published 1620)
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1-62). New York: Academic Press.
- Doyle, A. C. (1974). *The memoirs of Sherlock Holmes*. London: John Murray and Jonathan Cape. (Original work published 1893)
- Fazio, R. H., Sherman, S. J., & Herr, P. M. (1982). The feature-positive effect in the self-perception process: Does not doing matter as much as doing? *Journal of Personality and Social Psychology*, 42, 404-411.
- Fischhoff, B. (1975). Hindsight \neq foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 288-299.
- 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.
- Hand, L. (1951). Morals in public life. In *The spirit of liberty: Papers and addresses of Learned Hand* (pp. 225-252). New York: Knopf. (Collection published 1960)
- Henle, M. (1962). On the relation between logic and thinking. *Psychological Review*, 69, 366-378.
- Higgins, E. T., & Chaires, W. M. (1980). Accessibility of interrelational constructs: Implications for stimulus encoding and creativity. *Journal of Experimental Social Psychology*, 16, 348-361.
- Higgins, E. T., & King, G. (1981). Accessibility of social constructs: Information-processing consequences of individual and contextual variability. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality, cognition, and social interaction* (pp. 69-121). Hillsdale, NJ: Erlbaum.
- Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, 13, 141-154.
- Hovland, C. I., & Weiss, W. (1953). Transmission of information concerning concepts through positive and negative instances. *Journal of Experimental Psychology*, 45, 175-182.
- Jenkins, H. M., & Sainsbury, R. S. (1970). Discrimination learning with the distinctive feature on positive and negative trials. In D. Mostofsky (Ed.), *Attention: Contemporary theory and analysis* (pp. 239-273). New York: Appleton-Century-Crofts.
- Jennings, D. L., Amabile, T. M., & Ross, L. (1982). Informal covariation assessment: Data-based versus theory-based judgments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 211-230). Cambridge, England: Cambridge University Press.
- Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist*, 34, 107-117.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430-454.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 107-118.
- Lord, C. G., 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.
- Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Platt, J. R. (1964). Strong inference. *Science*, 146, 347-353.
- Popper, K. R. (1965). *The logic of scientific discovery*. New York: Harper & Row.
- Quattrone, G. A. (1982). Overattribution and unit formation: When behavior engulfs the person. *Journal of Personality and Social Psychology*, 42, 593-607.
- Ross, L., & Lepper, M. R. (1980). The perseverance of beliefs: Empirical and normative considerations. In R. A. Shweder (Ed.), *New directions for methodology of behavioral science: Fallible judgment in behavioral research* (pp. 17-36). San Francisco: Jossey-Bass.
- Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, 32, 880-892.
- Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 544-551.
- Smoke, K. L. (1933). Negative instances in concept learning. *Journal of Experimental Psychology*, 16, 583-588.
- Snyder, M. (1981). Seek, and ye shall find: Testing hypotheses about other people. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition. The Ontario Symposium* (Vol. 1, pp. 277-303). Hillsdale, NJ: Erlbaum.
- Snyder, M., & Campbell, B. (1980). Testing hypotheses about other people: The role of the hypothesis. *Personality and Social Psychology Bulletin*, 6, 421-426.
- Snyder, M., & Swann, W. B., Jr. (1978). Hypothesis-

- testing processes in social interaction. *Journal of Personality and Social Psychology*, 36, 1202-1212.
- Swann, W. B., Jr. (1984). Quest for accuracy in person perception: A matter of pragmatics. *Psychological Review*, 91, 457-477.
- Taylor, S. E., & Fiske, S. T. (1978). Salience, attention, and attribution: Top of the head phenomena. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 249-288). New York: Academic Press.
- Trope, Y., & Bassock, M. (1982). Confirmatory and diagnosing strategies in social information gathering. *Journal of Personality and Social Psychology*, 43, 22-34.
- Tuchman, B. W. (1984). *The march of folly From Troy to Vietnam*. New York: Knopf.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology* (pp. 135-151). Middlesex, England: Penquin.
- Wason, P. C. (1969). Regression in reasoning? *British Journal of Psychology*, 60, 471-580.
- Wason, P. C., & Golding, E. (1974). The language of inconsistency. *British Journal of Psychology*, 65, 537-546.
- Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning Structure and content*. London: Batsford.

Received October 10, 1983

Revision received August 6, 1984 ■

Acknowledgment

This journal relies heavily on the expert help and generous cooperation of a large number of reviewers. Once a year, in this issue of the journal, we normally list the people who have served in this capacity. Unfortunately, however, due to the recent editorial turnover and related changes in administrative personnel, our records for 1984 are incomplete. The Editor and Associate Editor would nonetheless like to take this occasion to thank the many persons who have given so generously of their time in helping to evaluate the manuscripts submitted during the past year to the section on Attitudes and Social Cognition.