The Bias Blind Spot: Perceptions of Bias in Self Versus Others

Emily Pronin
Daniel Y. Lin
Lee Ross
Stanford University

Three studies suggest that individuals see the existence and operation of cognitive and motivational biases much more in others than in themselves. Study 1 provides evidence from three surveys that people rate themselves as less subject to various biases than the "average American," classmates in a seminar, and fellow airport travelers. Data from the third survey further suggest that such claims arise from the interplay among availability biases and self-enhancement motives. Participants in one follow-up study who showed the better-than-average bias insisted that their self-assessments were accurate and objective even after reading a description of how they could have been affected by the relevant bias. Participants in a final study reported their peer's self-serving attributions regarding test performance to be biased but their own similarly self-serving attributions to be free of bias. The relevance of these phenomena to naïve realism and to conflict, misunderstanding, and dispute resolution is discussed.

Cognitive and social psychologists have documented a number of specific cognitive and motivational biases that compromise lay inference and judgment (for reviews, see Dawes, 1998; Gilovich, 1991; Nisbett & Ross, 1980; Plous, 1993; Tversky & Kahneman, 1974). Everyday observation confirms the existence of such biases. We find that our adversaries, and at times even our peers, see events and issues through the distorting prism of their political ideology, their particular individual or group history and interests, and their desire to see themselves in a positive light. When we reflect on our own views of the world, however, we generally detect little evidence of such bias. We have the impression that we see issues and events "objectively," as they are in "reality." We would concede, perhaps, that some of our views have been shaped by our unique personal experience or group identity, but we feel that in our own particular case these factors have led to increased insight rather than bias.

It is this perceived asymmetry in susceptibility to bias that provides the focus of the present article. We propose that people recognize the existence, and the impact, of most of the biases that social and cognitive psychologists have described over the past few decades. What they *lack* recognition of, we argue, is the role that those same biases play in governing their *own* judgments and inferences.

This proposal of an asymmetry in perceptions of bias arises from recent accounts of "naive realism" (Griffin & Ross, 1991; Pronin, Puccio, & Ross, 2001; Ross & Ward, 1996; also see Ichheiser, 1970), which hold that people think, or simply assume without giving the matter any thought at all, that their own take on the world enjoys particular authenticity and will be shared by other openminded perceivers and seekers of truth. As a consequence, evidence that others do not share their views, affective reactions, priorities regarding social ills, and so forth prompts them to search for some explanation, and the explanation most often arrived at, we argue, is that the other parties' views have been subject to some bias that keeps them from reacting as the situation demands. As a result of explaining such situations in terms of others' biases, while failing to recognize the role of similar biases in shaping their own perceptions and reactions, individuals are likely to conclude that they are somehow less

Authors' Note: Research reported in this article was supported in part by National Institute of Mental Health Grant MH-44321. Emily Pronin was supported under a National Science Foundation Graduate Research Fellowship during the conduct of this research. We gratefully thank Zachary Dillon for his assistance in carrying out Study 3. Please address correspondence concerning this article to Emily Pronin, Department of Psychology, 1444 William James Hall, Harvard University, Cambridge, MA 02138; e-mail: pronin@wjh.harvard.edu.

PSPB, Vol. 28 No. 3, March 2002 369-381 © 2002 by the Society for Personality and Social Psychology, Inc.

subject to biases than the people whom they observe and interact with in their everyday lives.

The prior literature documents the fact that social perceivers are well aware of specific biases that influence the responses of their peers. Studies have found that individuals readily cite, and even exaggerate, the role of biases such as the "fundamental attribution error" (Van Boven, Kamada, & Gilovich, 1999), self-serving attributions of personal responsibility (Kruger & Gilovich, 1999), and the tendency to rely on personal self-interest in making decisions "for the greater good" (Miller & Ratner, 1998) in accounting for others' responses. This evidence, however, does not explicitly address the question of an asymmetry in assessments of one's own versus others' susceptibility to bias.

Our present studies are designed to document and explore invidious distinctions people make between their own and others' susceptibility to bias. We further seek to demonstrate that whereas the "bias blind spot" can be seen as a particular instance of the so-called better-thanaverage effect, and can serve to enhance a generally positive view of self, it cannot be understood entirely in motivational terms. Specifically, we seek to demonstrate that the perceptual phenomena involved in naïve realism, and in asymmetries in cognitive availability (Tversky & Kahneman, 1973), play a role in creating this biased perception of freedom from bias. In this context, we note that there is ample evidence that people do rate themselves as better than average in a wide variety of domains such as sense of humor, morality, and driving ability (e.g., Dunning, Meyerowitz, & Holzberg, 1989). There is even evidence that people rate themselves as generally more objective than their peers (Armor, 1999). But we also note that in other domains ranging from joke-telling ability to computer programming prowess (two domains, we would contend, in which people who lack the relevant skill are likely to be all too aware of their deficiencies), people perceive themselves to be below average (Kruger, 1999). We anticipate, therefore, that our attempts to link asymmetric perceptions of bias to asymmetries in cognitive availability also will sharpen our understanding of the better-than-average effect.

Overview of the Present Studies

Our first study featured a series of three surveys, each of which described a number of specific biases. Participants were asked to assess their own susceptibility to these biases and that of individuals in some particular comparison group. In the second and third surveys, these assessments also were sought with regard to short-comings and biases that we thought individuals would be aware of at the time they were operating (or after the fact) and that we thus thought would not produce the postulated asymmetry in assessments of susceptibility.

Study 2 examined individuals' awareness that they had been subject to a specific bias (the better-than-average effect) immediately after they had offered the relevant self-ratings and then been explicitly prompted about the nature of that bias. In Study 3, participants first rated the validity of a (social intelligence) test on which they and a peer had just succeeded or failed and then assessed the degree to which their own and their peers' validity ratings had been subject to self-serving bias.

STUDY 1: PERCEIVED BIAS IN SELF AND OTHERS

Survey 1: Self Versus the "Average American"

In our first survey, participants were asked to indicate how much they, and the average American, showed eight specific biases that have been well documented in previous research.

METHOD

Participants. A group of 24 Stanford students enrolled in an upper-level psychology class (titled Misunderstanding, Conflict, and Dispute Resolution) completed the survey as homework. None of the biases dealt with in the survey had yet been discussed in the class, although participants might have learned about some of them in other psychology courses.

Procedure and questionnaire. Participants each received a booklet describing eight specific biases: self-serving attributions for success versus failure, dissonance reduction after free choice, the positive halo effect, biased assimilation of new information, reactive devaluation of proposals from one's negotiation counterparts, perceptions of hostile media bias toward one's group or cause, the fundamental attribution error (FAE) in "blaming the victim," and judgments about the "greater good" influenced by personal self-interest. The descriptions used the neutral term "effect" or "tendency" rather than the nonneutral term "bias." For example, self-serving attributional bias was described as follows:

Psychologists have claimed that people show a "self-serving" tendency in the way they view their academic or job performance. That is, they tend to take credit for success but deny responsibility for failure; they see their successes as the result of personal qualities, like drive or ability, but their failures as the result of external factors, like unreasonable work requirements or inadequate instruction.¹

Thirteen participants were asked first about their own susceptibility to each of the eight biases (i.e., "To what extent do you believe that *you* show this effect or tendency?") and only then about the susceptibility of the average American to each (i.e., "To what extent do you believe the average American shows this effect or ten-

dency?"), whereas the remainder (n=11) rated the average American before themselves. In addition to rating the average American, participants also rated the susceptibility of one of their parents. Ratings were made on 9-point scales anchored at 1 (not at all) and 9 (strongly), with the midpoint of 5 labeled somewhat. Items were counterbalanced with respect to the order in which the various biases were listed.²

RESULTS AND DISCUSSION

Examination of composite scores for the eight biases revealed that participants, as predicted, reported themselves less susceptible to the biases (M = 5.31) than the average American (M = 6.75), t(23) = 8.31, p < .0001. This asymmetry in ratings for self versus the average American was apparent (see Figure 1) and statistically significant (all ps < .02) for each individual bias. Of interest, participants also rated their *parent* as less susceptible to each bias than the average American (all ps < .01), although only in the case of the "positive halo effect" did they rate themselves differently (and in fact as more susceptible) than their parent.

Although these data are consistent with our "bias blind spot" hypothesis, an obvious alternative possibility exists. That is, as students in a prestigious university, participants in our survey might simply have been making an invidious comparison between their own analytical capacities (and presumably those of their classmates as well) and those of the hypothetical "average American." Our second survey was designed in part to rule out this possibility and in part to begin exploring our contentions regarding the mediating role of cognitive availability.

Survey 2: Self Versus the Average Fellow Classmate

This survey asked participants to rate their susceptibility to various biases relative to that of fellow students in a seminar course—a comparison target that was less hypothetical and more relevant to our participants than the average American. We also added three survey items intended to buttress our contention that the invidious comparison in question pertained not to all biases or shortcomings but only to those whose impact one is unaware of at the time.

METHOD

Participants. Thirty students enrolled in the following year's class of the same upper-level psychology seminar used for our first survey served as participants. (The survey was again presented as homework, prior to any inclass discussion of the relevant biases or shortcomings.)

Procedure and questionnaire. This survey featured two key modifications from the previous one. First, the comparison target specified was not the "average American" but instead the participant's classmates (i.e., the "average

Psych 182/256 student"). Second, this survey included three new items. Two of these items, procrastination and fear of public speaking, involved personal limitations that are not cognitive or motivational biases and are limitations that the individual is likely to be well aware of at the time he or she is putting off the dreaded task or experiencing pre-speech jitters. The third new item concerned the planning fallacy, a personal shortcoming that is a cognitive bias but one that our conceptual analysis suggested should not show the self-other asymmetry in assessments of susceptibility. (One who confidently expects to submit a manuscript on the first of September, promises collaborators "conservatively" to have it done by the first of October, only to find that it is still on the "to do" list in the waning days of November, is likely unaware, at the time of the initial estimate, that the relevant bias is operating. However, such awareness does dawn on one later—although, if the present authors' experiences are any indication, such after-the-fact acknowledgment does not seem to confer immunity to recurrences.⁴) As in the case of our previous survey, appropriately labeled 9point scales accompanied all items. (Again, order of response target and order of bias were varied between participants but, as before, no relevant order effects were found.)

RESULTS AND DISCUSSION

Our analyses confirmed that participants perceived themselves as less biased with respect to our original eight-item composite (M=5.05) than their average classmate (M=5.85), t(28) = 4.64, p<.0001 (see Figure 2), although the relevant asymmetry appeared less pronounced and consistent than in the earlier survey where the comparison target had been the average American. Examination of individual items revealed that the relevant difference in assessments of self versus others reached statistical significance only for four of the eight items (the self-serving bias, the "victim-blaming" FAE, reactive devaluation, and the positive halo effect).

More relevant to our conceptual analysis, the participants in our second survey did not rate themselves as less prone than their peers to the three personal limitations—procrastination, public speaking, and the planning fallacy—that one is likely to be aware of by the time they are done making their influence felt (see Figure 2). In fact, participants reported themselves to be somewhat more prone to these three limitations (M = 5.68) than their classmates (M = 5.05), although this "reversal" did not reach significance, t(28) = 1.49, p = .15. The relevant contrast between the means for the eight items on which an invidious self-other comparison was predicted and these three new items for which no difference was predicted was statistically significant, t(28) = 3.31, p < .003.

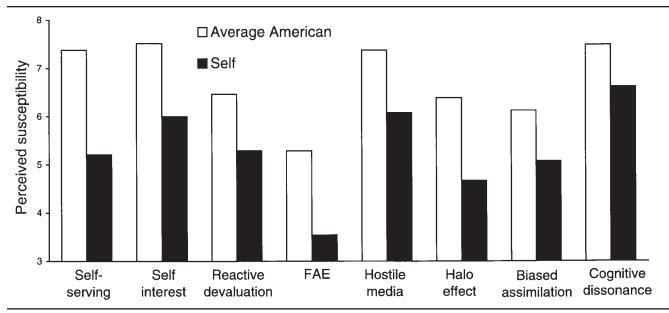


Figure 1 Participants' perceptions of their own and the "average American's" susceptibility to eight biases in judgment and inference (Survey 1). NOTE: FAE = fundamental attribution error.

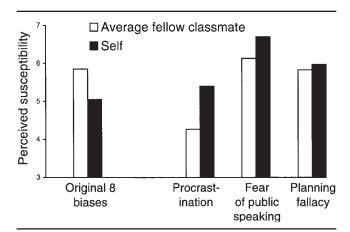


Figure 2 Participants' perceptions of their own, and their fellow classmates', susceptibility to the eight biases for which invidious self-other comparison was expected and three items for which such invidious comparison was not expected (Survey 2).

In summary, the results of our second survey replicated our prior results (albeit with somewhat smaller differences in susceptibility ratings for self vs. others) despite the change in comparison targets from the hypothetical average American to specific individuals who were both familiar to our participants and similar in status. Furthermore, data from the three new items offered preliminary support for our conceptual analysis. Participants did not serve their self-enhancement motive by claiming less susceptibility than their peers when the shortcoming or bias in question was one likely to be highly available either at the time that it was manifesting itself or afterward when its consequences became obvious. In these cases, partici-

pants saw themselves as being just as flawed as their peers, if not more so!

Although this second survey served to sharpen the message of our initial survey, two issues remained. First, would the relevant invidious distinction between self and others continue to be apparent when we examined the assessments of individuals who were older, less highly selected, less "elite," and perhaps more worldly and/or less immodest than our Stanford students? Second, could we disentangle the role of cognitive availability from that of social desirability in influencing individuals' perceptions of differential susceptibility to bias? To answer these questions, we generated a broader range of biases and took our survey instruments to a busy metropolitan airport.

Survey 3: Self Versus the Average San Francisco International Airport (SFO) Traveler

Our third survey featured both a more diverse population (travelers at SFO, the main airport serving the San Francisco Bay area) and a more diverse set of biases—that is, biases that we assumed would vary widely with respect to their social desirability and their cognitive availability to the individual affected by them. The seven new biases added to the 11 items employed in Survey 2 were as follows: friend enhancement (positive illusions about the capacities of one's friends), trust of strangers (overconfidence in the kindness and good intentions of strangers), trust of borrowers (unwarranted trust that borrowers will return items one has loaned them), generous attribution (attributing a person's charitable contributions to generosity rather than social pressure or

convenience), downward comparison bias (selectively comparing oneself to people worse off in times of illness or hardship), upward comparison bias (selectively comparing oneself to people who have successfully met the challenges one is currently facing), and gambler's fallacy (failure to appreciate the independence of individual wagering outcomes).

To investigate the role of social desirability and cognitive availability in producing the "bias blind spot," we next asked six raters—three graduate students and three advanced undergraduates in social psychology—to read the descriptions of each "tendency" and then to make two types of assessments about its status vis-à-vis potential mediating factors. Ratings of both availability (i.e., how aware does one ultimately become of having committed this tendency?) and desirability (i.e., how negatively does this tendency reflect on someone showing it?) showed acceptable interrater reliability (Cronbach's α = .77 and .87, respectively).

Using median splits for the relevant ratings, we thus established a set of eight low availability biases and eight high availability biases. (The two shortcomings in our survey that were not biases—that is, fear of public speaking and procrastination—were of course excluded from these classifications, but it is worth noting that, as anticipated, both rated high on this availability dimension.) We used a similar median split procedure to establish a set of eight high negative (i.e., socially undesirable) and eight low negative biases.

Inspection revealed that our lists of low availability and high "negativity" biases overlapped considerablyas, conversely, did our lists of high availability and low "negativity" biases. Specifically, seven of the eight biases from the initial survey were rated low in availability and high in negativity, with the exception being dissonance reduction after free choice, which was rated low in both availability and negativity. Similarly, the planning fallacy (added in Survey 2) and six of the eight biases added in Survey 3 were rated high in availability and low in negativity (with the exceptions being the gambler's fallacy, which was rated high in availability and high in negativity, and the generous attribution effect regarding donationgivers, which was rated low in availability and low in negativity). Furthermore, the positive correlation between undesirability and lack of availability for the 16 biases was high (r = .49, p = .05). This association between availability and social desirability would inevitably frustrate our attempts to disentangle the impact of these two factors on perceived susceptibility. But, we believe it ultimately helped us to appreciate the way in which these two factors combine to produce the perceived asymmetry in susceptibility assessments regarding self and others.

METHOD

Participants. While awaiting flights at SFO, 76 individuals of varying ages and ethnic backgrounds completed our survey, for which they received a free lottery ticket.

Procedure and questionnaire. The questionnaire format was similar to that of our previous surveys except for one important change that served to make the relevant self-other comparison more explicit. Whereas our previous surveys had asked participants to rate susceptibility for self and other on separate scales, this survey simply provided a description of each bias and asked first how widespread it was among that day's SFO travelers $(1 = not \ at \ all, 3 = slightly, 5 = somewhat, 7 = moderately, 9 = extremely)$ and then how susceptible the participant was to that bias "relative to the average SFOer" that day $(1 = much \ less \ than the average SFOer, 3 = slightly \ less, 5 = same, 7 = slightly more, 9 = much more than the average SFOer).$

To produce questionnaires brief enough for airport use, two versions of the instrument were prepared, such that about half of our participants (n = 39) were asked to assess themselves relative to their fellow travelers with regard to the positive halo effect, self-interest bias, "victimblaming" FAE, hostile media effect, planning fallacy, kind stranger effect, fear of public speaking, gambler's fallacy, and selective downward comparison, whereas the remainder were asked to assess themselves with regard to the self-serving bias, reactive devaluation, assimilation bias, dissonance reduction, trusting lender effect, charitable attribution of donations, procrastination, friend enhancement effect, and selective upward comparison. Questionnaires were counterbalanced with respect to the order of the various biases (and again, our analyses revealed no significant main effects or interactions involving this variable).

RESULTS AND DISCUSSION

The results of this survey generally replicated those of the two previously reported ones. The "explicit" comparison format adopted in Survey 3 revealed that participants claimed to be less biased than members of the relevant comparison group (in this case, fellow airport travelers)⁵ on the eight-item composite of biases employed in our two previous surveys, t(74) = 3.67, p < .0005, but not on procrastination, public speaking, or the planning fallacy (all ts < 1). It is also worth noting that our SFO participants rated the seven new biases introduced in this survey to be almost as pervasive among their fellow travelers (M = 5.66) as the eight biases from our original Survey 1 (M = 5.89), t(74) = 1.39, p = .17.

Our primary research questions, however, involved the association between asymmetric ratings of susceptibility on one hand and low availability and/or low social desirability on the other. As predicted, participants saw themselves as much less susceptible than other travelers to the eight biases that our raters judged to be low in cognitive availability (M = .53), t(74) = 3.51, p < .001, but equally susceptible to the eight biases that our raters judged to be high in such availability (M = .04, ns) (see Table 1). The difference between these two means, in turn, was highly significant, t(74) = 3.28, p < .002.

Our analyses revealed a similar association between perceptions of own versus others' susceptibility to a bias and the social desirability of that bias. As shown in Table 1, participants saw themselves as much less susceptible than their peers to the biases deemed by our raters to be low in social desirability (M = .61), t(74) = 3.99, p < .0002, but equally susceptible to the biases deemed to be high in social desirability (M = -.05, ns). Once again, the relevant difference between these two means was highly significant, t(74) = 4.25, p < .0001. These two factors of desirability and availability also proved to be associated with a "bias blind spot" when we simply correlated the degree of claimed relative personal "unsusceptibility" regarding the 16 biases with our raters' assessments of the cognitive availability (r = -.42, p = .08) and negativity (r=.70, p<.01) of those biases.⁶

Clearly, in the case of assessments regarding bias, low cognitive availability and low social desirability are associated both with each other and with the asymmetric assessments of susceptibility to bias that are the focus of our investigation. This confounding, which we noted in discussing the assessments of our raters, is more than an annoying methodological problem. Rather, it is a clue about the way the two factors combine and reinforce each other in producing the relevant phenomenon. Awareness that one is susceptible to a given bias (or for that matter, awareness of one's susceptibility to any shortcoming or difficulty) is apt to foster, even motivate, the assessment that such susceptibility is difficult to avoid, and not a sign of more general moral or intellectual deficiency. Conversely, the awareness that one is subject to a given shortcoming or bias that is particularly negative is likely to make one work hard to eliminate that shortcoming and to deny one's susceptibility to it. To some extent, one will continue to manifest that shortcoming to others only if and when one is unable to recognize that one is doing so.

In summary, the findings from Survey 3 continued to offer considerable support for our general contention about perceptions of bias in self versus others, and they suggest that a combination of—or even an interplay between—low cognitive availability and perceived negativity plays a role in creating and perpetuating the relevant invidious self-other distinction.

Our two concluding studies show that perceptions of personal superiority in resisting bias occur even in situa-

BLE 1: San Francisco International Airport (SFO) Travelers' Ratings of Their Susceptibility to Bias: Comparison of High Versus Low Cognitive Availability and High Versus Low Social Desirability Biases (Survey 3)

Composite of Biases	Perceived Freedom From Susceptibility to Bias	
	M	SD
Cognitive availability		
Low availability biases	.53*	1.32
High availability biases	.04	1.11
Social desirability		
Low desirability biases	.61*	1.34
High desirability biases	05	1.13

NOTE: Ratings are relative to the susceptibility of the "average SFOer." Mean ratings greater than 0 reflect the predicted tendency to rate oneself as less susceptible to the relevant biases than the average SFOer, ratings of 0 reflect the tendency to rate oneself as equally susceptible, and ratings less than 0 reflect the tendency to rate oneself as more susceptible.

tions where situational demands and constraints would seemingly make it easy to recognize and acknowledge personal susceptibility and difficult to claim immunity.

STUDY 2: DENYING PERSONAL SUSCEPTIBILITY TO THE BETTER THAN AVERAGE EFFECT

The three surveys conducted in Study 1 established that people report being less susceptible than their peers to various cognitive and motivational biases. This invidious self-other comparison arises, we suggest, because individuals generally are unaware of biasing influences that are being exerted on them when they are in the process of making judgments or inferences. Accordingly, they are inclined to infer bias not in their own judgments or inferences but in those that differ from their own. Study 2 further explored this proposed lack of awareness with respect to the so-called better-than-average effect (Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995; Dunning et al., 1989). Essentially, we first invited participants to make a series of self-assessments that we anticipated would produce this effect, and we then sought to determine whether participants would deny having been guilty of it, even after just exhibiting the bias and being given an explicit description of it.

Method

Participants. In exchange for introductory psychology course credit, 91 Stanford students responded to an omnibus questionnaire that included the two-page survey constituting the present study.

Procedure and questionnaire. Participants were first asked simply to rate themselves "relative to other

^{*}p<.001, based on two-tailed t tests comparing relevant mean with 0.

Stanford students" on six personality dimensions. Three of these dimensions were positive (dependability, objectivity, and consideration for others) and three were negative (snobbery, deceptiveness, and selfishness). Ratings were provided on 9-point scales anchored at 1 (*much less than the average Stanford student*) and 9 (*much more than the average Stanford student*), with the midpoint appropriately labeled *same as the average Stanford student*.

After providing these ratings, participants turned to the next page of their surveys, where they immediately found the following description of the better-than-average effect:

Studies have shown that on the whole, people show a "better than average" effect when assessing themselves relative to other members within their group. That is, 70-80% of individuals consistently rate themselves "better than average" on qualities that they perceive as positive, and conversely, evaluate themselves as having "less than average" amounts of characteristics they believe are negative.

Participants were then informed that "for purposes of our study, it would be useful to know the accuracy of your self-assessments on the previous page." To that end, they were asked to indicate how they thought they would be rated on the relevant dimensions by the "most accurate, valid, and objective resources available." Their response choices were as follows (excluding the bracketed labels, provided here for the benefit of our readers):

___ The objective measures would rate me *lower on positive characteristics* and *higher on negative characteristics* than I rated myself. [Recognize Bias]

___ The objective measures would rate me *neither more positively nor more negatively* than I rated myself. [Claim Objectivity]

___ The objective measures would rate me *higher on positive characteristics* and *lower on negative characteristics* than I rated myself. [Claim Modesty]

Results and Discussion

Overall, participants claimed both to possess more of the positive characteristics listed (M=6.44), t(78)=14.19, p<.001, and less of the negative ones (M=3.64), t(78)=10.94, p<.001, than the average Stanford student (designated by the midpoint of 5 on the relevant 9-point scales). Moreover, 79 of the 91 individual participants (87%) personally offered a mean rating across the six scales that reflected a claim of being better than the average of their peers.

More relevant to our immediate concerns, did participants recognize having been influenced by the better-than-average effect? When subsequently asked to evaluate their self-assessments in light of what they had read about this effect, only 19 of the 79 individuals (i.e., 24%)

claiming better-than-average status indicated that their responses had been biased. Thus, the large majority of participants who claimed better-than-average status insisted on this status. They either claimed that their initial self-ratings had been accurate and objective (63%) or that these ratings had actually been too modest (13%); that is, that they felt that objective measurement would show that they merited even better ratings than the already better-than-average ones they had previously given. Clearly, some participants had both succumbed to the relevant bias and then denied having done so, even after hearing an explicit description of it and being invited to acknowledge its influence, $\chi^2(1) = 9.97$, p < .002.7

These data suggest that even the immediate experience of having displayed a particular bias, and then being given an explicit description of it (one which presents the bias as a common human tendency), was insufficient to prompt confessions of susceptibility equal to that of one's peers. Neither the subjective experience associated with the occurrence of the bias, nor the description of the bias as common (and therefore presumably a forgivable foible rather than a loathsome transgression), nor an explicit invitation and context that further would have made it easy and socially desirable to acknowledge one's frailty, was sufficient to prompt a mea culpa from the majority of our participants.

Could it be that our participants recognized their susceptibility but simply were reticent or embarrassed about retracting their claims of superiority? Two considerations make this suggestion unconvincing. First, the instructions and description of the study provided to participants emphasized the experimenters' interest in knowing whether such a bias was present. Second, the participants were told that "70-80%" of people display the relevant bias. Social desirability and experimenter demand considerations thus should have made it easier for participants to accept the investigators' invitation to acknowledge their frailty than to continue to insist that they were better than the average of their classmates. Yet they resisted that invitation.

In Study 1, we argued that individuals' perceptions that they show less-than-average susceptibility to bias is not a simple case of self-enhancement but rather a phenomenon that arises at least in part because of availability considerations. Yet, in the present study, we have chosen to focus on the phenomenon most frequently cited as evidence of self-enhancement—that is, the better-than-average or "Lake Woebegone" effect—in pursuing our investigation of this bias blind spot. We will discuss this irony at some length in our general discussion. For now, we shall merely assert our contention that the better-than-average effect (like the bias blind spot itself) results from a complex interplay between motivational factors and cognitive ones, including availability or aware-

ness considerations. This contention regarding mediation may help to explain the apparent lack of a better-than-average effect in certain "high availability" domains, such as procrastination.

In summary, the results of Study 2 provide further evidence for our claim that people are poor "bias detectors" when they focus their attention on themselves. However, we have further claimed that people are rather good and sophisticated bias detectors (indeed, because of naïve realism, often overzealous detectors) when they focus their attention on their peers. Study 3 tested both claims simultaneously by having participants consider the possibility of bias both in assessments they had just made and in similar assessments that had just been made by one of their peers. This time, the relevant bias involved self-serving attributions regarding success versus failure.

STUDY 3: RECOGNIZING SELF-ENHANCEMENT BIAS IN SELF VERSUS OTHERS

Study 1 showed that individuals claim to be less susceptible than their peers to a variety of cognitive and motivational biases. Study 2 showed that people who have just manifested a particular bias insist that their assessments have been unbiased even under conditions that should make it easy to acknowledge such bias. Study 3 was designed to determine whether pairs of individuals who had just manifested a particular bias would recognize its impact on the assessments of their peer but not on their own assessments. In this study, participants first received either a high or low score on a putative test of social intelligence, then were given the opportunity to rate the validity of that test, and finally were asked to assess the possibility of self-enhancement or "ego-defensiveness" both in their own ratings of test validity and in those of a peer who ostensibly had experienced a test outcome opposite to their own. Our hypothesis, again, was that participants would see more bias in their peers' assessments than in their own.

Method

Participants. Twenty same-sex pairs of Stanford students (10 male pairs and 10 female pairs), none of whom had any prior acquaintanceship with their partner, participated in Study 3. The participants either received introductory psychology course credit or a \$7 payment for the half-hour study.

Procedure. After arriving at the laboratory and completing a written consent form, the two participants were introduced to each other. The experimenter, a male undergraduate, then presented them with a "social intelligence test," which he described as follows:

Our research is concerned with social intelligence and social sensitivity. In particular, we're looking at the valid-

ity of different types of measures of this ability. So today, we're going to give you a standard social intelligence task in which you'll be asked to match people's appearances with their own written self-descriptions. There are exactly 18 pictures and 18 descriptions. The descriptions state, in the individuals' own words, a description of their career, a hobby, and something they like. Your performance will be assessed by comparing the number of correct matches you provide with the performance of other Stanford students who have taken the test for the purposes of building standardized norms on it.

The test that participants were asked to undertake was one that would be absorbing and seemingly high in face validity. It also was designed to allow us to provide nonveridical but plausible feedback regarding performance (because there was no obvious strategy to be followed in completing the task and no way to know whether one had succeeded or failed either in making a given match or in outperforming one's peers). Because these characteristics could make the test appear somewhat difficult for participants, the experimenter also offered the following assurance before participants undertook their task:

One other thing—most people find that the test seems pretty hard—don't worry about it. We've found that social intelligence is actually very intuitive and determined by mostly unconscious processes, so don't worry if you don't feel like you're getting "the right answers" because you're probably doing fine.

When the participants had completed the "test" the experimenter collected their answer sheets and then returned a short time later with personalized score sheets indicating the participants' "scores" and "percentile ranks relative to other Stanford students." The participant in each pair randomly assigned to the success condition received a score of 14 out of 18 and an "80th percentile" ranking, whereas the participant assigned to the failure condition received a score of 6 out of 18 and a "30th percentile" ranking. This false feedback was followed by two questions inviting participants to share their impressions about the validity of the test. The first question was, "Do you think this task is a valid measure of social intelligence?" ($1 = not \ at \ all \ valid$, $7 = very \ valid$). The second question was, "Do you think individuals' performance on this task will be correlated with their performance on other more established measures of social intelligence (i.e., that those who score well on this test would also score well on other social intelligence tests and that those who score poorly would also score poorly on other social intelligence tests)?" (1 = scores would be not at all correlated, 7 = scores would be very highly correlated).

After collecting the participants' answers, the experimenter offered them some time to "discuss . . . [their] general thoughts about the test's validity" while he prepared "debriefing sheets." After returning, the experimenter thanked the participants for their cooperation and provided written information that began by indicating that we were "trying to determine the validity of various measures of social intelligence." They were reminded that it was important that we "get as good an idea as possible of how valid you feel this task is in indicating your social intelligence." Participants then found the following description of the self-serving bias:

Psychologists have claimed that some people show a "self-protective" tendency in the way they view their performance on any kind of test. That is, they tend to feel that tests on which they perform well are more valid than tests on which they do not perform as well.

The participants were reminded that the researchers needed them to be as careful as possible in assessing the validity of the social intelligence test at hand and were told that the researchers wanted to know "whether this bias may have affected your assessments and your partner's assessments." They were further assured that having the "most accurate data available" was crucial to us and were accordingly asked to consider both their partner's and their own earlier assessments of test validity. To this end, they were presented both with their own and their partner's scores and their own and their partner's prior evaluations of test validity. Dyads were counterbalanced with respect to order of item presentation, such that participants in half the pairs considered their partner's evaluations first (n = 20), whereas the other half of participants considered their own first. The first of the two items asked, "To what extent do you think your own [your partner's own] score influenced your [your partner's] evaluation of the test?" (1 = not at all influenced, 7 = very much influenced). The second item asked, "To what extent do you think your [your partner's] evaluation of the test reflects a 'self-protective' tendency?" (1 = not at)all reflects, 7 = very much reflects).

Finally, participants were debriefed about the real purpose of the experiment and were provided with appropriate assurances that other participants had similarly believed the test to be real and that their scores were not reflective of their actual social intelligence and had, in fact, been assigned randomly. Where any additional concerns were expressed, the experimenter offered a version of "process debriefing" (Ross, Lepper, & Hubbard, 1975) to make sure that no one left the laboratory with any erroneous notions about his or her social intelligence.

Results and Discussion

As expected, the large majority of our dyads (i.e., 15 of 20 dyads) displayed evidence of the relevant bias. That is, the individual in the pair who received success feedback perceived the test as a more "valid measure of social intelligence" (M = 4.53) than the individual who received failure feedback (M = 2.13), t(28) = 7.51, p < .0001. (The means, it is interesting to note, suggest that participants who received failure feedback checked a point farther from the midpoint of 4 on the scale than did those who received success feedback—an indication that "failing" participants were more adamant about the test's invalidity than "successful" participants were about its validity.) Further evidence of the self-serving bias was provided by ratings on the second item relating to test validity; that is, "successful" participants expected a stronger correlation between scores on this test and scores on "other more established measures of social intelligence" (M = 4.91) than did their "failing" partners (M = 2.82), t(20) = 4.43, p < .0003.

Most important, in terms of our present concerns, participants were more inclined to detect such bias in their partner's evaluation of the test's validity than in their own; that is, in the participant pairs who had displayed evidence of the self-serving bias, individuals generally thought that their partner's score had "influenced his/her evaluation of the test" (M=4.30) more than they thought their own score had influenced their own evaluation of the test (M = 3.57), t(28) = 2.12, p < .05. Of interest, for the measure of perceived bias that merely named the bias (rather than describing it), the effect was somewhat weaker. Participants reported that their partner's "evaluation of the test reflected a 'self-protective' tendency" (M = 4.07) more than did their own evaluation (M = 3.67), but this difference did not reach statistical significance, t(28) = 1. (No significant main effects or interactions involving gender of respondents or order in asking about bias in self versus other were found on either measure.)

We next compared assessments made by participants given high social intelligence scores (whose partners had received low scores) with assessments made by participants given low social intelligence scores (whose partners had received high scores). This comparison was prompted in part by our finding that high-scoring participants had been less adamant about the test's validity than low-scoring participants had been about its invalidity. Not surprisingly, in light of this difference, participants' perceptions of how little bias they had shown, relative to their partner, were much more pronounced among those who had "succeeded" (M = 3.67) while their partner had "failed" (M = 5.00), t(13) = 2.87, p < .02, than among those who had "failed" (M = 3.40) while their partner had "succeeded" (M = 3.60), t < 1. In other

words, the predicted invidious comparison effect was accompanied by an understandable tendency for participants to rate seemingly "ego-defensive" claims of low test validity as more reflective of bias than seemingly "self-enhancing" claims of high test validity.

The results of Study 3 supported our hypothesis by demonstrating that participants were more likely to see bias in others than in themselves under conditions in which both self and other seemingly were making self-serving assessments of test validity. Indeed, it highlighted the specific situation in which "winners" see bias in the complaints of "losers" and (to a lesser extent) that in which losers see bias in the fairness claims of winners.

Unfortunately, our design did not include a condition in which participants evaluated the degree of bias in responses that were identical to their own (e.g., claims of low test validity by a fellow "failing" participant or claims of high test validity by a fellow "succeeding" participant). But, regardless of whether people would make invidious distinctions in such cases, the present result is noteworthy. It suggests a process, anticipated by the "naïve realism" account that helped prompt the present research, by which people who disagree about matters of high hedonic relevance come to see each other as biased. In particular, such perceptions may be especially likely to arise when the disagreement occurs between people who have done well and think that the test of merit provided by the world has been valid or fair and people who have done poorly and think that the relevant test of merit has been invalid and unfair.

GENERAL DISCUSSION

The results of our three studies suggest that knowledge of particular biases in human judgment and inference, and the ability to recognize the impact of those biases on others, neither prevents one from succumbing nor makes one aware of having done so. Indeed, our research participants denied that their assessments of their personal qualities (Study 2) and their attributions for a particular success or failure (Study 3) had been biased even after having displayed the relevant biases and reading descriptions of them.

Three related questions are prompted by these research findings, each of which we have tried to address in the course of these studies: First, what accounts for the invidious distinction people make in assessing their own versus others' susceptibility to cognitive and motivational biases? In particular, is it cognitive availability, motivation to be seen in a positive light (by oneself or others), or some combination of these factors? Second, under what specific circumstances do people seem to recognize rather than overlook or deny their susceptibility to bias? Third, what is the relationship between the specific tendency for people to deny their susceptibility to particular

biases and their more general tendency to rate themselves as "above average" with respect to many personal attributes? A final question—regarding what light these findings may cast on understanding misunderstandings and conflicts between groups—also merits further attention.

Sources of the Invidious Self-Other Distinction

We suggested at the outset of this article that availability biases, coupled with the experience that others do not share some of our views, foster the impression of an "objective" self in a world of "biased" others. As earlier accounts of naïve realism suggest (Griffin & Ross, 1991; Ross & Ward, 1996), we hold our experience of people, objects, and events in our world to be veridical, more or less "unmediated," perceptions of reality. We further extend this epistemic stance to include perceptions of more complex objects such as evaluations of arguments, attributions of cause and effect, and even interpretations of historical fact. This sense that we perceive reality without any distortion arises in part because we lack direct access to the cognitive and motivational processes (to say nothing of the underlying biochemical processes) that influence our perceptions (Nisbett & Wilson, 1977; see also Nisbett & Ross, 1980, chap. 9). Because we lack immediate access to these processes, we do not consciously experience or otherwise enjoy direct access to their biasing effects.

Instead, the operation of bias must be inferred. Such inferences are precisely those we make when there seems to be a discrepancy between what another individual perceives, or at least what he or she claims to perceive, and what we assume to be reality. Because our peers, and especially our adversaries, often fail to share our views, we inevitably infer that they are less objective than we are. Furthermore, we readily apply what we know about specific biases from observing our peers and from the wisdom handed down to us by our sages to diagnose specific failures on their part to see the world "as it is." What we are slow to recognize, of course, is that our views of the world are no less subject to those same specific biases. Again, our conviction at the time we are making a specific judgment is that we are merely "seeing things as they are" and then "calling them as we see them."

This line of reasoning helps us to reconcile two seemingly contradictory theoretical assertions entailed in "naïve realism." The first, an assimilation hypothesis, is that ceteris paribus people tend to assume that others share their views (Ross, Greene, & House, 1977). The second, a differentiation hypothesis, is that people often feel that their adversaries hold views that are extremely different from—and opposing to—their own (e.g., Robinson, Keltner, Ward, & Ross, 1995). Our contention is that in most cases people do indeed expect that others will

share their views. But, when these others fail to do so, people are likely to see those with whom they disagree as unreasonable and as unable to view things in an objective manner. The reason for this conclusion, again, is that individuals have faith in the "realism" or objectivity of their own views, and are thus likely to assume bias on the part of those who fail to share those views. And it is this tendency to view others as influenced by bias that leads individuals to the conclusion that their opponents hold extreme and dogmatic points of view.

In considering the factors underlying this tendency to see others as the ones who are susceptible to bias, we tried to disentangle cognitive availability considerations from social desirability considerations by querying our participants about biases that differed along these two dimensions. But we failed in this disentangling attempt, largely because none of the biases assessed by our raters as highly available also were assessed to be highly negative. We do not think that this "confounding" reflected a failure in our methodology. Rather, we think that awareness that one is susceptible to a given shortcoming makes one try to eliminate it and, after failing to do so, to assume it is both ubiquitous and forgivable. As a result, we conclude that cognitive and motivational factors reinforce each other in producing the illusion that one is less susceptible to bias than one's peers.

Recognition of One's Own Biases and Other Personal Limitations

The conceptual analysis offered in this article anticipates the types of biases and other personal shortcomings to which individuals do see themselves as susceptible—in some cases as even more susceptible than their peers. The key factor, we argue again, is cognitive availability. When we suppress nagging doubts in vouching for a friend who seems to have consistently been unlucky in his employment history, in lending money to a relative, or in promising to meet a deadline (or when, after the fact, we recognize such folly) we do come to realize that we have succumbed to the relevant type of "overconfidence" or "wishful thinking." More mundane failings such as daydreaming or procrastinating similarly may be highly salient to the "guilty" party and thus not likely to prompt invidious self-other comparisons.

Indeed, consideration of the role of availability in these social comparisons suggests a number of personal shortcomings to which people should assume themselves to be more prone than their peers. Such shortcomings likely involve private thoughts or feelings that people normally do not share with each other. Self-doubts and fears (including fears about public speaking and other kinds of performance anxiety), "perverse" fantasies, and envy of friends provide obvious examples of private events that people do not normally discuss

with their peers. As a consequence, these private experiences may give rise to erroneous feelings of personal deficiency or deviance.

Reconsidering the Better-Than-Average Effect

There seem to be a number of specific domains in which people rate themselves to be no less flawed than their peers. What implications do such exceptions have for our understanding of the much-documented betterthan-average effect? We have no doubt that people are generally motivated to see themselves positively relative to their peers and to adopt self-evaluation criteria that further that goal (Dunning et al., 1989). But, we think that people also may change their behavior to justify greater self-regard. We further contend that people's motivation to see themselves positively may sometimes be thwarted by availability biases and other factors that make them feel worse about themselves than would be objectively justified, or deemed reasonable by peers. Our own view, following Dunning et al. (1989) and Kunda (1987), is that people do want to think well of themselves and generally succeed in doing so. But assimilation biases, cognitive and perceptual availability, biased sampling of information, and other nonmotivational factors influence people's views of themselves, just as they influence the assessments they make about other entities and objects. In short, students of inference and judgment should recognize motivational influences, but such influences do not merit some sovereign status in our attempts to explain human biases or shortcomings.

Implications for Social Misunderstanding and Conflict

Displays of cognitive and motivational bias are inevitable products of the way we all see and understand the world. Perceptions and accusations of bias in others, coupled with denial of bias in self, are similarly inevitable. Misunderstanding, mistrust, escalation of conflict, and unwarranted pessimism about the ability to find common ground with those with whom we disagree become likely consequences when we attribute disagreements and bias not to ordinary psychological processes but to evil strategic designs or the unique traits of our "opponents." This unhappy scenario is suggested by some recent data we have been collecting in the volatile political climate of Northern Ireland. In surveying partisans immersed in the struggle between Nationalists and Unionists regarding the "Good Friday Agreement," we are finding that people perceive their opponents as more susceptible than their own side to various psychological biases (being ruled by emotions rather than reason, distorting the nature of past events in a manner that serves the interest of one's own side, etc.) that create barriers to resolution of the conflict.

This phenomenon is all too familiar to individuals who have worked in the field to create more constructive interethnic dialogues via "second track" diplomacy or public peace processes. In the best of all possible worlds, people would come to recognize their own biases and to recognize that they are no less susceptible to such biases than their adversaries. In the imperfect world in which we live, people should at least endeavor to practice a measure of attributional charity. They should assume that the "other side" is just as honest as they are (but not more honest) in describing their true sentiments—however much these may be distorted by defensiveness, self-interest, propaganda, or unique historical experience.

Adversaries also would be well advised to engage the efforts of sophisticated third parties. They should do so, we hasten to add, not in the hope that those third parties will offer a uniquely objective perspective on the truth, for such a perspective would likely be dismissed by both partisan groups as biased against them (Vallone, Ross, & Lepper, 1985; see also Morris & Su, 1999). Rather, they should do so in the hope that the third parties will help them in the search for common ground and for a future that both sides would find better than the status quo.

NOTES

- 1. Verbatim descriptions of all biases used in our research are available on request.
- 2. No effects involving order of question were found, either with respect to order of rating self versus the "average American" or order of the various biases listed in the questionnaire. Accordingly, these variables receive no further attention in the present report.
- 3. This finding apparently reflected our participants' specific reluctance to claim less susceptibility to bias than a family member, not—as will become apparent from the data to be reported in our subsequent survey—any general tendency (see Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995) to rate real individuals more charitably than hypothetical ones.
- 4. One additional small change was made to our materials (for this survey and for Survey 3). Rather than beginning each bias description with the phrase, "Psychologists have claimed that people", the descriptions instead began, "Psychologists have claimed that some people" This one-word addition removed the demand to report, regardless of one's actual perceptions, that one's peers were highly susceptible and perhaps also the demand to report similar susceptibility for self.
- 5. In these analyses, we simply compared participants' responses on each scale to the midpoint of the scale (which indicated equal susceptibility of self and the average San Francisco International Airport traveler ["average SFOer"]).
- 6. We initially made several attempts to disentangle these effects through statistical analysis but ultimately rejected this strategy. We were obliged to do so because there was only one bias (the tendency to take charitable donations at face value, without recognizing possible situational factors or less-than-noble motives on the part of the donor) that was rated below the median in cognitive availability and above the median in social desirability and only one bias (the gambler's fallacy) that ranked below the median in social desirability and above the median in availability. Accordingly, we chose to focus our attention on the possible reasons for the relevant linkages between these two factors and for their association with perceived susceptibility rather than

attempt to generate a longer list of "atypical" biases that people thought very negative and yet were aware of displaying.

7. In calculating this χ^2 , we assumed that 46 students (the closest whole number to that representing half of the 91 respondents) could objectively be classified as "better than average." Thus, of the 79 students who rated themselves as better than average, at most 46 (i.e., 58% of this 79) could have been accurate in claiming that status, even if none of the respondents in the study meriting such a designation had modestly declined to make such a claim.

REFERENCES

- Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psy*chology, 68, 804-825.
- Armor, D. A. (1999). The illusion of objectivity: A bias in the perception of freedom from bias. *Dissertation Abstracts International: Section B*, 59, 5163.
- Dawes, R. M. (1998). Behavioral decision making and judgment. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 497-548). Boston: McGraw-Hill.
- Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving appraisals of ability. *Journal of Personality and Social Psy*chology, 57, 1082-1090.
- Gilovich, T. (1991). How we know what isn't so: The fallibility of human reasoning in everyday life. New York: Free Press.
- Griffin, D., & Ross, L. (1991). Subjective construal, social inference, and human misunderstanding. In M. P. Zanna (Ed.), Advances in experimental social psychology (Vol. 24, pp. 319-359). San Diego, CA: Academic Press.
- Ichheiser, G. (1970). Appearances and realities. San Francisco: Jossey-Bass.
- Kruger, J. (1999). Lake Woebegone be gone! The "below-average effect" and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology*, 77, 221-232.
- Kruger, J., & Gilovich, T. (1999). "Naive cynicism" in everyday theories of responsibility assessment: On biased assumptions of bias. Journal of Personality and Social Psychology, 76, 743-753.
- Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psy*chology, 53, 636-647.
- Miller, D. T., & Ratner, R. K. (1998). The disparity between the actual and assumed power of self-interest. *Journal of Personality and Social Psychology*, 74, 53-62.
- Morris, M. W., & Su, S. K. (1999). The hostile mediator phenomenon: Egocentric standards of fairness lead disputants to see mediators as favoring the opponent. Unpublished manuscript, Stanford University.
- Nisbett, R. E., & Ross, L. (1980). Human inference: Strategies and short-comings of social judgment. Englewood Cliffs, NJ: Prentice Hall.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Plous, S. (1993). The psychology of judgment and decision making. New York: McGraw-Hill.
- Pronin, E., Puccio, C. T., & Ross, L. (2002). Understanding misunderstanding: Social psychological perspectives. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, UK: Cambridge University Press.
- Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "Naive realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology*, 68, 404-417.
- Ross, L., Greene, D., & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13, 279-301.
- 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.

Ross, L., & Ward, A. (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. In T. Brown, E. Reed, & E. Turiel (Eds.), Values and knowledge (pp. 103-135). Hillsdale, NJ: Lawrence Erlbaum.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207-232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.

Vallone, R. P., Ross, L., & Lepper, M. R. (1985). The hostile media phenomenon: Biased perception and perceptions of media bias in coverage of the Beirut massacre. *Journal of Personality and Social Psychology*, 49, 577-585.

Van Boven, L., Kamada, A., & Gilovich, T. (1999). The perceiver as perceived: Everyday intuitions about the correspondence bias. *Journal of Personality and Social Psychology*, 77, 1188-1199.

Received October 18, 2000 Revision accepted June 5, 2001