Objectivity in the Eye of the Beholder: Divergent Perceptions of Bias in Self Versus Others

Article *in* Psychological Review · July 2004

Impact Factor: 7.97 · DOI: 10.1037/0033-295X.111.3.781 · Source: PubMed

3 authors, including:

Thomas Gilovich
Cornell University
112 PUBLICATIONS **8,293** CITATIONS
SEE PROFILE

Lee Ross
Stanford University
95 PUBLICATIONS **15,982** CITATIONS
SEE PROFILE

Psychological Review
2004, Vol. 111, No. 3, 781–799

0033-295X/04/$12.00 DOI: 10.1037/0033-295X.111.3.781

THEORETICAL NOTE

Objectivity in the Eye of the Beholder: Divergent Perceptions of Bias in Self Versus Others

Emily Pronin
Princeton University

Thomas Gilovich
Cornell University

Lee Ross
Stanford University

Important asymmetries between self-perception and social perception arise from the simple fact that other people's actions, judgments, and priorities sometimes differ from one's own. This leads people not only to make more dispositional inferences about others than about themselves (E. E. Jones & R. E. Nisbett, 1972) but also to see others as more susceptible to a host of cognitive and motivational biases. Although this *blind spot* regarding one's own biases may serve familiar self-enhancement motives, it is also a product of the phenomenological stance of *naive realism.* It is exacerbated, furthermore, by people's tendency to attach greater credence to their own introspections about potential influences on judgment and behavior than they attach to similar introspections by others. The authors review evidence, new and old, of this asymmetry and its underlying causes and discuss its relation to other psychological phenomena and to interpersonal and intergroup conflict.

> And why beholdest thou the mote that is in thy brother's eye but considerest not the beam that is in thine own eye?
>
> —Matthew 7:3 (King James Version)

This familiar biblical quotation describes an age-old double standard in the way people perceive themselves versus their peers. We suspect that people not only are subject to this double standard but also are inclined to believe that their peers are more subject to it than they are themselves. In the present article, we argue that people readily detect or infer a wide variety of biases in others while denying such biases in themselves. We place this argument in the larger context of theory and research on the relationship between self-perception and social perception. In particular, the ideas we advance can be seen as an extension of Jones and Nisbett's (1972) conceptual analysis of divergent actor–observer attributions, with the focus of our analysis shifting from judgments about traits to judgments about biases.

Our analysis begins with the observation that attributions about others often are prompted by evidence that their response to a given situation or issue differs from one's own. We then relate this observation to the broader epistemic stance of *naive realism* (Ross & Ward, 1995, 1996), the defining feature of which is the conviction that one sees and responds to the world objectively, or "as it is," and that others therefore will see it and respond to it differently only to the extent that their behavior is a reflection of something other than that reality. One attributional possibility is that the "divergent" response in question reflects the others' personality traits or dispositions. The other possibility, which has particularly important consequences when it comes to contentious social and political issues, is that the others' judgments and decisions reflect cognitive or motivational biases that distort reality.[1]

Naive realism, we suggest, thus gives rise to the conviction that others in general, and others who disagree with us in particular, are

Emily Pronin, Department of Psychology, Princeton University; Thomas Gilovich, Department of Psychology, Cornell University; Lee Ross, Department of Psychology, Stanford University.

Research reported in this article was supported in part by National Institute of Mental Health Grants MH44321 and MH63524 and National Science Foundation Grant SES0091416. We thank Daryl Bem, Leaf van Boven, Nicholas Epley, Chip Heath, Joachim Krueger, Justin Kruger, and Timothy Wilson for their helpful comments concerning this article.

Correspondence concerning this article should be addressed to Emily Pronin, Department of Psychology, Princeton University, Green Hall, Princeton, NJ 08544. E-mail: epronin@princeton.edu

[1] Definitions of bias in the literature are many, varied, and subject to controversy. Most refer to influences that lead to judgments or decisions that prove erroneous or suboptimal by some objective criterion or that violate some well-accepted axiom or normative criterion. Many analysts, however, have defended either the rationality or the functionality of judgmental tendencies or strategies that other psychologists cite as illustrative of nonnormativeness or suboptimality (see Gigerenzer, 1996; Gilovich & Griffin, 2002). Alternative, more subjective definitions of bias have emphasized the individual's willingness to characterize the influence in question as unwanted and undesirable (e.g., Wilson & Brekke, 1994). Given our present concerns, which focus not on the nature of bias but on perception of self–other differences in susceptibility, we have no need to resolve these definitional issues. When we refer to biases, we are simply referring to sources of influence that are commonly regarded (and certainly regarded by the individuals in our studies) as biases because they meet at least one of these criteria.

more susceptible to biasing influences than we are ourselves (Pronin, Lin, & Ross, 2002). This conviction, in turn, is reinforced by people's tendency to treat their introspections about the basis of their own judgments and decisions as highly probative or "sovereign" but not to give similar weight to the introspections of others.

We begin by reviewing existing evidence about the perceived impact of a number of specific biases on self versus others. We also present evidence from new studies designed to show the generality of this self–other asymmetry and its particular applicability to attributions about political judgments. We then link this blind spot regarding one's own biases to the specific nonmotivational mechanisms we have proposed, and rebut the argument that it merely constitutes a special case of self-enhancement or ego defensiveness. Finally, we discuss the role that this *bias blind spot* and other related phenomena—such as the tendency to view the media as hostile to one's own positions (Vallone, Ross, & Lepper, 1985) and the tendency to be overly optimistic about the verdicts of disinterested third parties (Babcock & Loewenstein, 1997)—play in promoting interpersonal misunderstanding and intergroup enmity.

Asymmetric Perceptions and Attributions Regarding Self Versus Others

The relationship between the self and social perception has been a topic of longstanding interest to social scientists, starting with Cooley's (1902) account of the "looking glass self." Bem's (1972) suggestion that self-perception generally follows the same inferential rules as social perception was very similar in spirit and implications. It is interesting to note that Bem's "self-perception" account added the caveat that this similarity leads to parallel inferences about self and other in the face of similar behavioral information "[only] to the extent that internal cues are weak, ambiguous, or uninterpretable" (Bem, 1972, p. 2). However, Bem's empirical demonstrations served to question the role of such "internal cues" or other private information in actors' inferences about their own attitudes and dispositions. More specifically, those demonstrations showed that observers or "simulators" who knew only what overt response a given actor had made, and the situation in which it was made, offered the same attributions about actors that actors made about themselves. Nisbett and Wilson's classic (1977a) article extended Bem's argument by demonstrating that the access actors enjoyed to their private mental life did not prevent them from making causal inferences (some correct and some incorrect) similar to those made by observers who possessed the same outcome information and considered the same array of potential causes (see also Nisbett & Ross, 1980, chap. 9).

The best known account of systematic *divergence* in the assessments of actors versus observers was offered by Jones and Nisbett (1972). They argued that observers typically offer more dispositional (and less situational) attributions in accounting for actors' responses than actors offer themselves. In analyzing this difference, Jones and Nisbett claimed that although it may in some cases reflect the "actor's need to justify blameworthy action," it "may also reflect a variety of other factors having nothing to do with the maintenance of self-esteem" (1972, p. 80).

One factor they noted, and for which they offered convincing anecdotal evidence, involved actor–observer differences in the amount of information likely to be available about cross-situational variability in responses. But the factor they emphasized most, and the one for which they cited empirical evidence, was a difference in perspective or focus of attention. The actor's attention, they pointed out, is apt to be focused on the situational features he or she is monitoring and addressing, whereas the observer's attention (as Heider suggested in 1958) is focused on the behaving actor. Jones and Nisbett (1972) duly noted the apparent tension between their account of actor versus observer attributions and that offered by Bem (1972) and then proceeded to present data to document the existence of those attributional differences and to demonstrate the causal role played by differences in perspective.

It is worth emphasizing that there is no real contradiction between Bem's (1972) claims and those of Jones and Nisbett (1972), or between the evidence cited by Jones and Nisbett and that cited by Nisbett and Wilson (1977a). The apparent contradiction vanishes once one recognizes that Bem, like Nisbett and Wilson, was discussing instances in which actors and observers are attending to and relying on similar information, whereas Jones and Nisbett were dealing with cases in which actors and observers are attending to and relying on different information (see Nisbett & Ross, 1980, chap. 9, for a further discussion of these informational differences).

One source of "different information" mentioned by Jones and Nisbett involves private mental events. They noted that "typically, the actor has more, and more precise, information than the observer about his own emotional states and his intentions" (Jones & Nisbett, 1972, p. 85). Whereas the actor's knowledge of his or her own intentions is "direct," the observer's knowledge of those intentions is "indirect, usually quite inferior, and highly subject to error" (Jones & Nisbett, 1972, p. 84). We shall discuss this source of different information, and the possibility of differential weighing of it by actors versus observers, when we turn our attention from attributions about situational versus dispositional determinants of behavior to attributions about the existence and impact of bias.

Jones and Nisbett mention another source of divergence in the attributions of actors and observers, one that we shall also discuss at length. Specifically, they predict:

> The extent to which each actor behaves *differently* in the same situation should cause each to attribute the other's actions to internal dispositional factors. If actor A is attuned to the reality of the situation and sees himself as behaving accordingly, any variations in B's behavior will be attributed to B's *idiosyncratic interpretations of that reality*. (Jones & Nisbett, 1972, p. 88; italics added)

This passage, and the brief quotations cited just before about differences in access to private emotional states and intentions, constitute the extent of their consideration of these two sources of divergent attributions. In our analysis, we essentially take up the story of the differing attributions of actors and observers by expanding on these two factors.

Before we continue our analysis, and note its departure from that of Jones and Nisbett (1972), it is worth commenting on a study they described that is relevant to the second of these two factors (Nisbett, Caputo, Legant, & Marecek, 1973, Study 2). This study showed that actors made more dispositional attributions about their peers' selections of girlfriends and academic majors than they did about their own selections. Although we suspect that many of the

study participants probably made *different* selections of majors from their peers and almost certainly made different selections of girlfriends, the relevance of such differences was not considered or explored. In the present article, we proceed from the important fact that in everyday social interaction and in everyday encounters with media reports of current events, people often are confronted with instances in which others respond to issues and events differently from the way they themselves are responding. Indeed, it is this very discrepancy in response—especially when it violates one's expectation that reasonable people will respond similarly—that motivates much of the attributional work people do (Pyszczynski & Greenberg, 1981; Wong & Weiner, 1981).

Two attribution questions are thus intertwined: First, "Why did person X or group X not respond the same way I responded, or at least the way I think I would have responded?" Second, "Why did person X or group X not respond as I had expected they would respond in light of the demands and constraints of the situation at hand and the merits of the available response options?" The answer to both questions, we suggest, is apt to involve attributions about the person or group—attributions not only about the other's underlying traits but also about the other's views and priorities and about the normative versus nonnormative factors that are presumed to be dictating those views and priorities. This attributional process lies at the heart of Ross and Ward's (1995, 1996) account of naive realism, to which we now turn.

The Epistemic Stance of Naive Realism

> We tend to resolve our perplexity arising out of the experience that other people see the world differently than we see it ourselves by declaring that those others, in consequence of some basic intellectual and moral defect, are unable to see the things "as they really are" and to react to them "in a normal way." We thus imply, of course, that things are in fact as we see them and that our ways are the normal ways. (Ichheiser, 1949, p. 39)

The core of naive realism is the conviction that one perceives objects and events "as they are"—in other words, that there is an objective reality to which one's perceptions correspond in a more or less one-to-one manner (Ross & Ward, 1995, 1996). This conviction is inescapable and deep, and it governs our day-to-day functioning despite what we may know about the constructive nature of perception. Thus, even when we have learned that the colors and objects we perceive reflect the interaction between a world of molecules and energy sources and our particular human sensory processes, we continue to respond to the relevant stimuli in accord with our naive conviction about objective reality. We accept the word of scientists that this world is perceived very differently by other creatures—that the sounds, sights, and smells they experience may scarcely resemble our own. But we tend to regard those creatures as having an altered perspective on the objective properties of the reality we perceive, rather than vice versa.

Other aspects of naive realism follow directly from this basic tenet. Notably, we expect other reasonable and attentive people to perceive the same "reality" we do. Insofar as our concern is with the experience of physical objects and their properties, this expectation may cause few difficulties beyond an occasional disagreement about whether the thermostat is set too high or too low, whether the shade of yellow in the kitchen is too intense, or whether the catch was made with the wide receiver's feet inside or outside the field of play. But naive realism makes its influence felt not only in convictions about physical reality but also in convictions about complex social events and political issues. We cannot fully escape the conviction that we likewise perceive such events and issues as they "really are," and that other reasonable people who have the same information about those events and issues will, or at least should, perceive them similarly.

The final aspect of naive realism, and the one that speaks most directly to the focus of this article, involves our response to the fact that other individuals often do *not* respond similarly. In such cases, we cannot attribute their responses to the nature of the events or issues that elicited them because we deem our own *different* responses to be the ones dictated by the objective nature of those events or issues. Instead, as Ichheiser (1949) articulated, we infer that the source of their responses must be something about *them.* The inference may be, as Jones and Nisbett (1972) indicated, that the relevant response reveals something about the actor's personal traits or dispositions (selfishness or generosity, aggressiveness or passivity, etc.). Or, the inference may be that the response reveals something about the biased "lenses" through which the actor views the object, issue, or event in question. Such inferences of bias may be especially likely to arise in cases involving responses to social or political issues, where the opposing partisans can cite a host of possibilities ranging from self-interest and peer-group pressures to dissonance reduction and media brainwashing to explain the "error" in their adversaries' views.

Before reviewing evidence pertaining to this asymmetry in perceived bias in the self and others, we turn to a second determinant of this divergence. In particular, we examine the role played by the actors' introspections about their goals and motives, and the weight they assign to the reported goals and motives of others.

The Introspection Illusion: Actor Versus Observer Assessments of Phenomenological Reports

Most of us are willing, at least on occasion, to entertain the possibility that our own judgments or decisions are tainted by bias. "Is Jane's warmth and enthusiasm clouding my judgment about her plan for reorganizing our sales strategy?" "Am I finding tonight's *Nightline* guest more persuasive than I would otherwise because he is a prominent leader of my religion?" However, when we entertain such possibilities of bias, we are unlikely to find any phenomenological trace of the bias in question.

Indeed, introspection typically reassures us that the judgments in question were correct and supported by sound reasons. "Her estimates about prospective sales seem a bit optimistic, but she knows her own personnel and the market better than I do, and she has more to lose than any of us if the sales fall short." "Reverend Weems didn't offer much evidence for his argument, but his honesty and moral conviction shone through."

On the basis of such introspective evidence, we are apt to conclude that although we may well have been guilty of particular biases on some past occasions, we are innocent of bias in the specific assessment about which we have introspected. We do not claim to be immune to wishful thinking, overconfidence, defensiveness, closed mindedness, and a host of other inferential and judgmental failings; we just do not recognize that we are succumb-

ing to them in any particular assessment we are currently making (or else, of course, we would change that assessment until we felt that it was in accord with reality).

At the same time, we are not particularly comforted when others assure us that they have looked into their own hearts and minds and concluded that they have been fair and objective. In other words, we tend to treat our own introspections as something of a gold standard in assessing why we have responded in a particular manner and whether our judgments have been tainted by bias. By contrast, we treat the introspections of other actors as merely another source of plausible hypotheses—to be accepted or rejected as a function of their plausibility in light of what we know about the particular actor and about human behavior in general.

We refer to this asymmetry as the introspection *illusion* because the faith people have in the validity of their own introspections is misplaced. Although people can report accurately on the *contents* of their thoughts and deliberations, the psychological *processes* and the true determinants of their behavior are often inaccessible to introspection (Nisbett & Wilson, 1977a, 1977b). A considerable body of research provides evidence that the cognitive processes implicated in instances of biased reasoning or decision making are typically nonconscious (Dawson, Gilovich, & Regan, 2002; Ditto & Lopez, 1992; Dunning, Meyerowitz, & Holzberg, 1989; Kunda, 1987). Wilson and Brekke (1994) have argued that the difficulty of avoiding bias is exacerbated precisely because the relevant influence so often occurs outside of conscious awareness. Indeed, in the case of dissonance reduction and other types of motivated biases, it has been argued that the very effectiveness of such processes may *depend* on their occurring outside of awareness (e.g., Festinger, 1957; Lieberman, Ochsner, Gilbert, & Schacter, 2001).

Most people, furthermore, can recall instances in which they have accepted evidence that was at variance with their prior views and that they would have preferred to reject. In fact, introspection often produces the honest conviction that one acted as one did in spite of, not because of, one's private sentiments. Wilson (2002, p. 113) suggests, more generally, that it is the very *amount* of "inside information" that we possess that gives rise to misplaced confidence in its authority and authenticity as a guide in assessing the causes of our own behavior.

In making attributions about the causes of another person's actions or judgments, we obviously do not have access to corresponding amounts of inside information, including information about the presence or absence of introspective traces of bias. And, as noted earlier, even when that other person assures us that he or she finds no such traces, we are apt to be skeptical about the probative value of such an assurance. Instead, we are inclined to rely on our general theories of bias and the particulars of the situation at hand, giving weight to "circumstantial evidence" and the relative plausibility of the rival explanations or attributions in light of that evidence. In other words, we are apt to bring to bear exactly the sorts of abstract rules, principles, or schemata that attribution theorists have described (Jones & Davis, 1965; Kelley, 1967). Accordingly, when others' judgments seem highly correlated with their preexisting beliefs, we assume that they have been guilty of assimilating the evidence at hand to those beliefs. When their judgments seem highly correlated with obvious motives, we assume that the latter have played a role in dictating the former.

Evidence of Asymmetries in Perception of Bias

Over the past several decades, psychologists have documented the ways in which people, as Bruner (1957) famously observed, "go beyond the information given" in disambiguating stimuli. The types of responses investigated have ranged from the basic and visual–perceptual to the complex and social, and both cognitive and motivational influences have been explored at length. The key finding of this research is that expectations and contextual factors combine with needs and motives to determine, and at times distort, what people notice and how they interpret what they see and hear. As research on various sources of bias in everyday judgment developed, it was soon followed by research examining the extent to which individuals are aware of different sources of bias.

Although we are not aware of any published research that has addressed whether people are generally inclined to detect bias more in others than in themselves, numerous investigators have examined their research participants' awareness of their own susceptibility to particular biases. Often these examinations were incidental to the investigators' main purpose of demonstrating their participants' susceptibility to a given bias or documenting some consequence of that susceptibility. There have also been a number of studies in which the researchers sought to show that their participants would err systematically in estimating the impact of some biasing factor on their peers. Finally, several studies have directly compared assessments made about the self with assessments made about others vis-à-vis some specific bias.

In Table 1 we offer a sampling of all three types of studies. It would take us beyond the purposes of this article to describe each of these studies in detail. Instead we merely attempt here to convey the range of phenomena explored and the findings reported.

Studies of Perception of Specific Biases in Self or in Others

Studies calling for individuals to assess possible biasing influences on their *own* judgments and decisions have generally documented a failure to recognize such influence. Among the biases addressed have been the *hindsight bias* (Fischhoff, 1975), the *halo effect* (Nisbett & Wilson, 1977b; Wetzel, Wilson, & Kort, 1981), *rationalization* (Gilbert, Brown, Pinel, & Wilson, 2000), the *planning fallacy* (Buehler, Griffin, & Ross, 1994), and *intergroup bias* (Dovidio & Gaertner, 1991). These studies have mainly been concerned with documenting the nonconscious or implicit nature of the specific processes under investigation.[2] Nevertheless, the findings suggest that the proposed asymmetry in recognition of bias is likely to derive, at least in part, from individuals' "blindness" to their own susceptibility.

A number of studies have demonstrated that individuals do recognize, at least implicitly, the role that certain biasing influences may exert on other people (Krueger, 1998; Krueger &

[2] Not included in this review are the many intriguing studies involving *nonconscious priming* (see Bargh & Chartrand, 1999, for a review), in which participants are asked during debriefing whether they are aware of anything that may have influenced their behaviors or assessments. What makes the results of these priming studies noteworthy, in terms of our present concerns, is that the relevant influence occurs in the absence of any awareness of that influence on the part of the participants.

Zeiger, 1993; A. G. Miller, Baer, & Schonberg, 1979; Van Boven, Kamada, & Gilovich, 1999). Again, however, the investigators' primary concern in these studies was their participants' sensitivity to (or even exaggeration of) one particular source of bias on the part of their peers; the investigators were not addressing the broader issue of sensitivity to motivational, perceptual, or cognitive biases in general—much less an asymmetry in sensitivity to bias in others versus self.

Typically, the methodology used to demonstrate that people are aware of biases on the part of others has been indirect. For example, in demonstrating people's appreciation of the *correspondence bias* (Gilbert & Malone, 1995; Jones, 1990) or *fundamental attribution error* (Ross, 1977), Van Boven at al. (1999) never mentioned the bias per se. Instead, participants first gave a speech advocating a position that had been assigned to them, and then they were asked to predict how another individual would assess their actual position on the issue. The participants' predictions made it clear that they assumed the individual in question would show the familiar correspondence bias—in fact, comparison of these predictions with the actual responses of the individuals who rated the participants' attitudes made it clear that the participants tended to *overestimate* the extent of the correspondence bias.

Studies of Bias in Self Versus Others

To our knowledge, only one set of studies, reported in Armor's (1999) dissertation, has addressed the question of how people perceive their own versus others' susceptibility across a number of different biases. These studies called for participants to rate themselves relative to their peers on a single scale, rather than for separate self-assessments and assessments of others. When using such a comparative scale, people, on average, claimed their judgments to be "more objective" than their peers' judgments—a result that can be seen as an interesting instance of the pervasive *better-than-average effect* (e.g., Dunning, Meyerowitz, & Holzberg, 1989).

There are a small number of studies that have looked at perceptions of biasing influences on self and other separately (in which the motivation to see oneself as "less biased than average" may be dampened somewhat). With the notable exception of studies by Wilson, Houston, Etling, and Brekke (1996) on perceptions of the impact of "anchors" on own versus others' judgments, most have dealt with biases involving self-enhancement (or in-group enhancement) or self-interest (Epley & Dunning, 2000; Friedrich, 1996; Goethals, 1986; Heath, 1999; Kruger & Gilovich, 1999; D. T. Miller & Ratner, 1998; Van Boven, Dunning, & Loewenstein, 2000; Vivian & Berkowitz, 1992). These studies were generally designed to show that people overestimate the extent to which they personally are influenced by "objective" concerns and/or overestimate the extent to which others are influenced by "self-serving" concerns—that is, that people detect or infer self-enhancement and self-interest in others' judgments and behaviors at the same time that they are unaware of the influence of such processes on their own judgments and behaviors.

Epley and Dunning (2000), for instance, asked participants to predict their own behavior and that of an average student in specific situations where one's behavior might be influenced by self-interest (e.g., buying few vs. many flowers in support of an on-campus charity drive). Their participants' predictions suggested that they believed self-interest would play little role in determining their own decision to support the charity but that it would play a considerable role in determining others' decisions. These predictions turned out to be generally correct with regard to their peers but incorrect with regard to themselves. That is, the participants' own responses, like those of their fellow participants, were substantially influenced by self-interest.

An influential series of studies by D. T. Miller and Ratner (1998) on the "myth of self-interest" suggests that people sometimes overestimate the role that incentives and other self-interested motives play in the responses of their peers. Their participants, for example, overestimated the extent to which prospective blood donors would be influenced by economic incentives, and the extent to which smokers and nonsmokers would differ about the legitimacy of particular smoking restrictions. Moreover, they made such predictions even as they claimed that their *own* responses would be untainted by such motivational biases. Heath (1999) provided similar evidence, showing that workers assume others' on-the-job motivation is rooted primarily in extrinsic financial incentives, while reporting that they themselves are primarily motivated by intrinsic incentives such as the opportunity to learn new skills.

The Bias Blind Spot (Pronin et al., 2002)

Our contentions about bias blindness and its sources are not restricted to perceptions of self-serving motives. The theoretical mechanisms we have emphasized—that is, observations of differences between one's own responses and the responses of others, and the differing weight given to own versus others' introspections—should make their influence felt across a wide range of biases. What evidence is there of such generality?

To address this question, Pronin et al. (2002) examined a wide range of well-known biases, some motivational in character, some nonmotivational, and some reflecting both motivational and nonmotivational influences. In one study, Stanford undergraduates were asked to complete a questionnaire indicating how much they, and how much the "average American," showed a variety of specific biases that have been well documented in previous research. The list of biases included (a) self-serving attributions for success versus failure; (b) dissonance reduction after free choice; (c) the positive halo effect; (d) biased assimilation of new information to preexisting beliefs; (e) reactive devaluation of proposals from one's negotiation counterparts; (f) perceptions of hostile media bias toward one's group or cause; (g) the fundamental attribution error manifested by "blaming the victim" for outcomes largely determined by situational forces and constraints; and (h) allowing one's judgments about the "greater good" to be influenced by self-interest.

Each bias was described in a couple of simple sentences. For example, the description of the self-serving attributional bias read 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

Table 1
Prior Research on Perceptions of the Presence of Bias in Self and in Others

Domain	Specific bias studied	Participants' task/methodology	Principal findings
		Perceived bias in self	
Self-assessment	Better-than-average effect (Ehrlinger, Gilovich, & Ross, 2004)	Provide ratings of self (relative to average person) on several traits, then assess objectivity vs. self-enhancing nature of those ratings.	Ps rate self to be "better than average," then explicitly claim that their ratings were objective rather than self-enhancing.
Person perception	Halo effect (Nisbett & Wilson, 1977b)	Rate appearance, accent, mannerisms (and likability) of likable ("warm") vs. unlikable ("cold") instructor, then infer impact either of those 3 qualities on ratings of his likability or of his likability on those 3 ratings.	Ps fail to recognize "halo effect" in their assessments, instead assume instructor's appearance/accent/mannerisms impacted their liking for him (rather than vice versa).
	Halo effect (Wetzel, Wilson, & Kort, 1981)	Same ratings task as in Nisbett and Wilson (1977b), however some Ps told to "introspect" while forming relevant impressions (others *not* told to do so).	Nisbett and Wilson result replicated, with introspection neither increasing recognition nor decreasing commission of relevant bias.
Intergroup perception	Out-group bias (Dovidio & Gaertner, 1991)	Rate whether various positive and negative traits can "describe a person" after subliminal priming with sketch of white or black person; also complete self-report measure of racial attitudes.	White Ps show more positive *implicit* attitudes toward whites than blacks (i.e., faster response latency to positive traits after white prime); this occurs even among Ps claiming low prejudice on the *explicit* self-report measure.
Judgment/decision-making	Hindsight bias (Fischhoff, 1975)	Make likelihood estimate with benefit of "hindsight" information, then assess necessity of such information for one's accurate estimate.	Ps influenced by hindsight but think their estimates would have been very similar in the absence of such information.
	Planning fallacy (Buehler, Griffin, & Ross, 1994)	Pick a school project with an impending deadline, then predict date by which one will complete it, then indicate confidence level associated with that prediction.	Ps underestimate time they will need to complete tasks (i.e., the "planning fallacy") but are highly confident in their predictions.
	Rationalization/dissonance reduction (Gilbert et al., 2000)	Ps induced to reduce dissonance by increasing their liking for a (randomly assigned) outcome, then asked to account for relevant assignment.	Ps show no awareness of their dissonance reduction, instead assume a benevolent other was responsible for relevant assignment.
		Perceived bias in others	
Self-assessment	Self-enhancement bias (Krueger, 1998)	Predict degree to which another P will claim various traits and how positively he or she will rate those traits (predictions and actual responses of other are compared).	Ps correctly perceive that the other Ps self-enhance via idiosyncratic trait evaluation.
Person perception	Correspondence bias (Van Boven, Kamada, & Gilovich, 1999)	Deliver speech supporting assigned position to observer aware of this assignment, then predict inference that the observer will make about one's actual position on the issue.	Ps overestimate degree to which observer will make a correspondent inference.
	Correspondence bias (A. G. Miller, Baer, & Schonberg, 1979)	Write an essay supporting assigned position, then indicate own position and predict position that a reader of that essay (informed of that assignment) will infer one holds.	Ps predict that observers will make a correspondent inference.
	False consensus effect (Krueger & Zeiger, 1993)	View a fictitious P's preferences and predict his or her estimates of the percentage of other people that will have those preferences.	Ps predict that others' consensus estimates will be biased by those others' *own* preferences.

		Perceived bias in self vs. others	
Self-assessment	Better-than-average effect (Friedrich, 1996)	Learn about the bias by reading or hearing description of it, then assess susceptibility of self or of "average person."	Ps rate themselves as less susceptible to this bias than they rate the average person.
	Self-serving attributions (Kruger & Gilovich, 1999)	Allocate responsibility for a joint product, and also predict allocations that will be made by one's partner and by a neutral third party.	Ps predict that others will be "self-serving" in their allocations and that a neutral party will agree with them more than with their partner.
	Self-enhancement bias (Armor, 1999)	Assess centrality of different features to the definition of a positive trait and also how much one possesses those features, then rate objectivity of one's responses relative to other introductory psychology students (on single scale).	Ps evaluate features they possess as more central to positive traits but report being more objective than others in their responses.
	Self-interest bias (D. T. Miller & Ratner, 1998)	Predict influence of personal self-interest (e.g., financial incentives) on the decisions self and others would make (e.g., about whether to donate blood), then indicate what one's decision would be.	Ps overestimate impact of self-interest on decisions others would make while denying its impact on their own decisions.
	Self-interest bias (Heath, 1999)	Indicate extent to which one's own and others' on-the-job motivation is influenced by extrinsic factors (e.g., pay) vs. intrinsic ones (e.g., accomplishing something "worthwhile").	Ps perceive influence of extrinsic factors to be greater in the case of others' motivations than in the case of their own.
	Self-interest bias (Epley & Dunning, 2000)	Predict own and others' behavioral choices in circumstances where one might be influenced by self-interest (e.g., moves in a prisoner's dilemma game or decisions to donate money to charity); later have the opportunity to make actual choices.	Ps accurately predict that others' behavioral choices will reflect self-interest but fail to anticipate impact of self-interest on own choices.
	Self-interest bias (Goethals, 1986)	Predict own and others' behavioral choices in circumstances where those choices might be influenced by self-interest.	Ps predict that others' choices will be biased by self-interest but that their own choices will not.
Intergroup perception	In-group bias (Vivian & Berkowitz, 1992)	Evaluate work produced by in-group vs. out-group and predict evaluations by out-group members and by neutral third parties.	Ps predict out-group members will show bias in favor of their own group and that third-party evaluations will resemble their own.
Judgment/decision-making	Anchoring effect (Wilson, Houston, Etling, & Brekke, 1996)	Make frequency estimate (e.g., number of physicians in phonebook) after exposure to irrelevant numerical anchor, then assess impact of anchor on own vs. others' estimates (some Ps also "forewarned" about possible impact of anchor).	Ps fail to recognize impact of anchor on own estimates (even when forewarned) but claim that it will exert an influence on other people's estimates.
	Endowment effect (Van Boven, Dunning, & Loewenstein, 2000)	Assign value to item as a buyer (or seller) in a transaction, estimate value assigned by other party and by neutral third party not in transaction, and then assess whether endowment effect or "greed" influenced other.	Ps fail to recognize impact of endowment effect both on self and on others, but they assume that others will be biased by personal greed and that a neutral party will share their valuations.
	Conjunction rule violation (Armor, 1999)	Respond to a classic conjunction fallacy problem (see Tversky & Kahneman, 1983), rate relative objectivity of own response.	Ps show the relevant conjunction fallacy in their responses but report being more "objective" than others.
	Base rate neglect (Armor, 1999)	Respond to a classic base-rate neglect problem (Tversky & Kahneman, 1973), rate relative objectivity of own response.	Ps show the relevant base-rate neglect bias but report being more "objective" than others.
	Overconfidence effect (Armor, 1999)	Answer trivia questions by offering range within which one is 90% confident of accuracy, rate relative objectivity of own response.	Ps show the relevant overconfidence effect but report being more "objective" than others.

Note. P = participant.

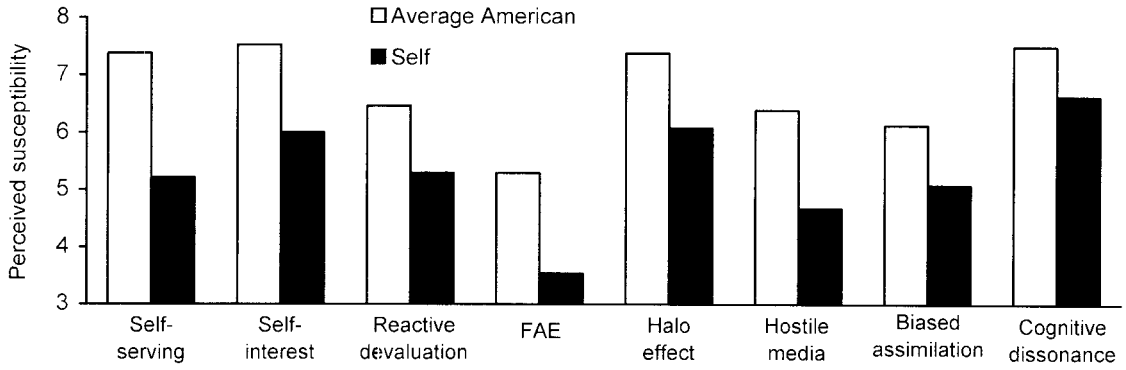

Figure 1. Perceptions of one's own versus others' susceptibility to eight different biases in human judgment and inference. FAE = fundamental attribution error. Adapted from "The Bias Blind Spot: Perceptions of Bias in Self Versus Others," by E. Pronin, D. Y. Lin, and L. Ross, 2002, *Personality and Social Psychology Bulletin, 28*(3), p. 372. Copyright 2002 by The Society for Personality and Social Psychology. Adapted with permission of Sage Publications.

drive or ability, but their failures as the result of external factors, like unreasonable work requirements or inadequate instruction.[3]

Consistent with our hypothesis, participants overwhelmingly reported that they personally were less susceptible to each of these biases than the average American (see Figure 1). Two additional surveys, one asking undergraduates to compare themselves with fellow students in a seminar and one asking San Francisco airport travelers to compare themselves with others traveling from the same airport on the same day, showed that this self–other asymmetry was not limited to students in a prestigious university making invidious comparisons between their own analytical capacities and those of the "average American." In all three surveys, the participants claimed to be less susceptible to the relevant biases than members of the stipulated comparison group.

Evidence suggesting that people tend to be oblivious to biases in themselves while readily recognizing the impact of bias on others is, of course, consistent with mounting evidence that people are motivated to view themselves, their assessments, and their outcomes in a positive light (e.g., Baumeister, 1998; Steele, 1988; Taylor & Brown, 1988). Accordingly, a motivational account of the bias blind spot, or at least an account that gives some weight to motivational considerations, seems reasonable and appropriate. Indeed, the pejorative connotation of the word *bias* is motivation enough for people to deny or understate, and perhaps even prevent themselves from perceiving, such bias in their own judgments and decisions. By contrast, no such motivation exists to deny bias in others, except in cases where those others' connections to us, high opinion of us, or support of our views gives us a reason to do so.

Our present contention, therefore, is not that motivational factors play no role in claims of relative freedom from bias, but rather it is that such claims do not arise *entirely* from such motives. Biased cognitive searches that provide after-the-fact justifications for actions or assessments that, on a priori grounds, seem tainted by bias may indeed be triggered by the desire to justify one's actions and defend oneself against the charge of self-interested bias. But searches that yield welcome justifications may also be distorted by availability or sampling biases that are not motivated. And although such biased searches may blind us to our shortcomings and enhance our sense of rationality in a way that is undeniably ego enhancing, the same nonmotivational biases can also produce unwarranted self-doubt and self-criticism.

This is the case, for example, when personal flaws—such as the tendency to daydream, procrastinate, or feel pangs of jealousy—are more apparent to the introspecting actor than to peers who are not privy to the actor's introspections (D. T. Miller & McFarland, 1987; Savitsky & Gilovich, 2003). In such cases, although the individual might be motivated to deny the negative self-assessment, the motive in question may fail to overcome the impact of the availability bias; in fact, availability considerations may lead to the erroneous judgment that one is especially susceptible to the relevant failing. In the next section of our article, we provide evidence for the operation of nonmotivational factors in producing the bias blind spot. In so doing, we begin to address an issue that motivational accounts of the relevant asymmetry in perceptions of bias are ill suited to tackle: that is, *when* we are most likely to see biases in others and *how* we remain oblivious to the same biases in ourselves.

Evidence Linking the Bias Blind Spot to Proposed Mechanisms

We have argued that the asymmetry in bias detection reflects the operation of two nonmotivational factors: the attempt to account for discrepancies between self and other, and the weight given to own versus others' introspections. We now present evidence linking these factors to people's perceptions of bias in others and to their convictions about their own freedom from such bias.

The Role of Perceived Attitude and Behavior Discrepancy in Attributions of Bias

The account of naive realism offered earlier suggests that inferences about bias in others follow directly from the discovery that those others do not share one's views about issues and events, and

[3] Verbatim descriptions of each bias are available from the authors.

from the attributions placed on such differences. In a sense, the problem involves a failure to distinguish one's subjective interpretations or construals of the relevant objects of judgment from the objects themselves. More specifically, people assume that while their own assessments reflect a logical, bottom-up progression from evidence and rational considerations to reasonable inferences and conclusions, others' assessments reflect a top-down process whereby preexisting motives and beliefs bias subsequent inferences and perceptions.

One line of relevant research suggests that people expect their peers to be less objective, and in a sense less fair, than they are themselves in situations that involve self-interest. These studies have shown that participants assume not only that others will fail to share their judgments in such situations, but that the relevant discrepancy in responses will be due to others' tendency to behave in a self-serving (or "group-serving") manner when it comes to allocating responsibility (Kruger & Gilovich, 1999) or assessing group performance (Vivian & Berkowitz, 1992). Consistent with the tenets of naive realism outlined earlier, participants in these studies further expect that "neutral" third parties will share *their* views about allocation of responsibility and quality of performance rather than the views of the other, "biased" party.

If individuals see superior objectivity in themselves in cases where they are simply anticipating disagreement, what happens when such disagreement is an established fact? One study reported by Pronin et al. (2002) addressed this question. Pairs of participants completed a putative test of "social intelligence" (the validity of which was supposedly under investigation), after which they received false feedback about their performance such that one individual was led to believe he or she had performed quite well and the other that he or she had performed poorly. With their own scores in hand, participants were each asked to assess the validity of the test. Later, during what they were told was a debriefing session, participants received a written description of the self-serving bias as it applied to attributions of test quality, and they then were asked, in light of their own test scores and evaluations of test validity and those of their fellow participant, to indicate the extent to which they believed the self-serving bias might have distorted their own and their partner's evaluations.

The results supported our current contentions. Participants were more likely to believe that their partner had provided a biased evaluation than that they themselves had. Thus, students who received high scores claimed that the test was reasonably valid and that their unsuccessful partner's claims to the contrary reflected ego-defensive bias. Conversely, those who received low scores were inclined to pronounce the test invalid and to suggest that their successful partner's claims to the contrary were a self-serving exercise in ego enhancement.

If attributions of bias are born in perceptions of disagreement, it follows that we should view people with whom we disagree as more biased than people whose views approximate our own. Support for this proposition was first provided in research by Robinson, Keltner, Ward, and Ross (1995). The main thrust of this research involved the demonstration that people are apt to overestimate the degree of polarization in their peers' views about contentious social and political issues such as abortion and affirmative action (and underestimate the degree to which even partisans' views are held with some ambivalence and appreciation of valid arguments of the "other side"). Most relevant to our present concerns, however, was the finding that partisans on both sides of a given issue saw their own side's views as more reflective of objective evidence and rational concerns, and less influenced by political ideology, than the other side's views.

More direct evidence linking perceptions of bias to perceptions of disagreement comes from a study in which respondents were asked (in the wake of the attacks of September 11, 2001) to specify which of five positions on U.S. policy toward Afghanistan best captured their own views on the issue (Ross, Pronin, & Lin, 2004). The five positions ranged from advocacy that the United States "launch an all out war against terrorism wherever it exists and threatens U.S. people and interests" to the suggestion that our country should "rely entirely upon diplomacy, admit our past mistakes and change U.S. policies that create resentment and poverty." Half of the participants then were asked to indicate the extent to which various factors had influenced their own position, the position they found "next most acceptable," and the position they found "least acceptable." The other half were asked to indicate the extent to which the relevant factors had influenced others who held the same position as themselves (but not the extent to which those factors had influenced them personally), as well the extent to which those factors had influenced the holder of the "next most acceptable" and "least acceptable" alternatives.

Four of the factors listed involved objective or normative considerations (e.g., knowledge of the culture and conditions of the area, use of logic and common sense), and four involved nonnormative considerations or biases (e.g., wishful thinking, ideological rigidity). The results of this study were as predicted (see Figure 2). Participants who were asked to account for their own views responded similarly to participants who were asked to account for the views of those advocating the same position as themselves. Both groups indicated that the views in question were strongly influenced by normative factors and only weakly influenced by nonnormative factors. In other words, participants assumed that peers whose views mirrored their own had arrived at those views through a process that was just as objective and attuned to the realities of the situation at hand as the process they had followed themselves. By contrast, the participants attributed support of the position that was "next most acceptable" more to biases and less to normative factors than they attributed their support of their own position. And, of course, they attributed the position they deemed least acceptable much more to biases and much less to normative factors than their own position.

Another study explored this link between perceptions of disagreement and attributions of bias with greater precision and across a broader range of issues (Ross, McGuire, & Minson, 2004). Stanford undergraduates first completed a one-page survey asking them to indicate their attitudes on 10 different issues, ranging from capital punishment and abortion rights to U.S. antiterrorism measures and the merits of various public figures, from Vice President Dick Cheney to Senator Hillary Clinton. Participants' anonymous responses were then collected and redistributed for rating such that each participant assessed a peer's responses without knowing the identity of that peer. As in the survey study just described, participants were asked to indicate the extent to which their own positions on the 10 issues and the positions of their assigned peer reflected various normative or objective considerations and various nonnormative considerations or biases. They were also asked to specify the degree of similarity between

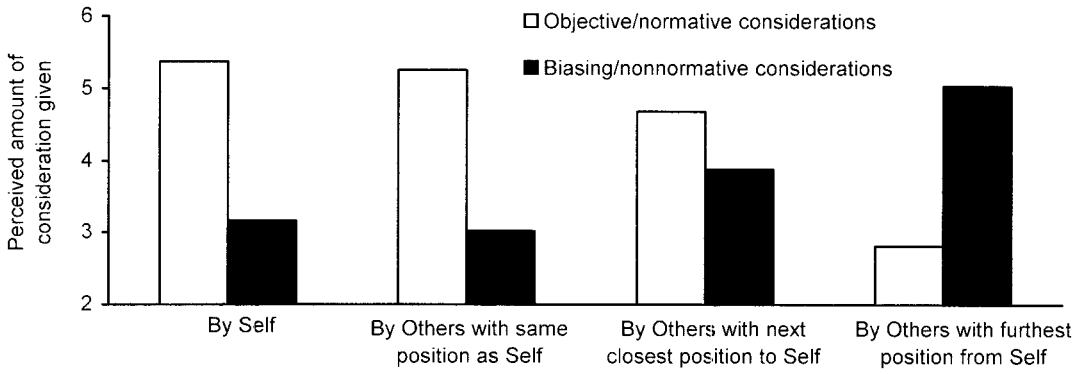

Figure 2. Estimate of weight given by self versus others to various normative and nonnormative considerations regarding views on U.S. intervention in Afghanistan after September 11, 2001 (from Ross, Pronin, & Lin, 2004).

their own views on the 10 issues and the views of the peer whose responses they had read.

As predicted, people thought their own attitudes reflected normative considerations to a significantly greater degree, and nonnormative considerations to a significantly lesser degree, than their peer's attitudes (see Figure 3). More important to our contentions about underlying mechanisms, there was a highly significant correlation between perceived discrepancy in own versus peer's attitudes and the degree to which the participants assumed that their

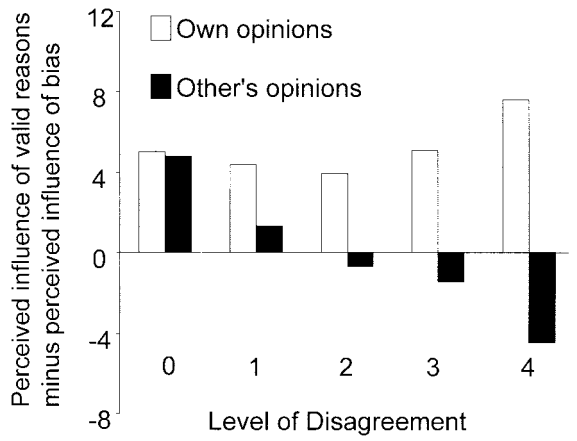

Figure 3. Perceptions of how much one's own attitudes and opinions, and those of a particular fellow student, are rooted in various normative ("valid") and nonnormative ("biasing") considerations (from Ross, McGuire, & Minson, 2004). Note that the levels of disagreement represented along the abscissa correspond to participants' self-rated similarity between their own position and that of the fellow student whose responses they had seen, with "0" representing perfect similarity. The data in this figure reflect that there was a high negative correlation between perceived discrepancy and other's good reasons minus biases ($r = -.47$) and a more modest positive correlation between perceived discrepancy and own bias minus own good reasons ($r = .26$). Examination of the separate components of other's bias minus other's good reasons further clarifies the relevant findings. Greater perceived discrepancy in views is correlated with the relevant difference measure vis-à-vis others less because of increases in willingness to infer bias vis-à-vis the other ($r = .18$) than because of decreases in willingness to infer good reasons ($r = -.58$).

peers had been influenced more by nonnormative than by normative factors.

At the risk of complicating our argument somewhat, it is worth noting that the distinction between a normative consideration and a biasing one can be subtle and subject to differences in construal. "Attention to the Mass Media" can be a normative factor when it is construed to mean careful attention to CSPAN or *Meet the Press* interviews with world leaders, but it can be a nonnormative factor or bias when it is construed to mean uncritical acceptance of the utterances of media pundits and the propaganda spread by partisans. Assessments of the influence of personal experience or group identity can be similarly problematic. People may be quite willing to acknowledge that their views have been affected by their status as a devout Christian, a member of an ethnic minority, or even a CEO of a Fortune 500 company. Such acknowledgment, however, is apt to be accompanied by the insistence that, in their own case, this status (and the experiences it afforded) has been uniquely enlightening—indeed, that it is the *lack* of such enlightenment that is making those on the other side of the issue take their misguided position. However, when it comes to assessing the views of others, especially those with whom we disagree, their special status and unique experiences are apt to be deemed a source of bias (albeit sometimes one that is viewed with sympathy and understanding rather than condemnation).

Ehrlinger, Gilovich, and Ross (2004) obtained evidence for just such an asymmetry. Their study examined the views of three pairs of complementary groups: (a) varsity athletes and intramural athletes, (b) ethnic minority students and non-ethnic-minority students, and (c) Arabs and Jews. Members of these groups were asked to rate how much a person's membership in a particular group (i.e., the group to which they belonged or the complementary group) affects that person's ability to reason clearly about an issue relevant to the group. As anticipated, participants consistently asserted that their own group memberships had more of an enlightening effect, and less of a biasing effect, than memberships in the complementary group (see Figure 4).

Our claim here, we recognize, is dangerously close to being tautological. If one were aware that a given influence was compromising the accuracy of one's present judgments and knew what a more accurate judgment would be, one presumably would modify the judgment in question and whatever decisions follow from it.

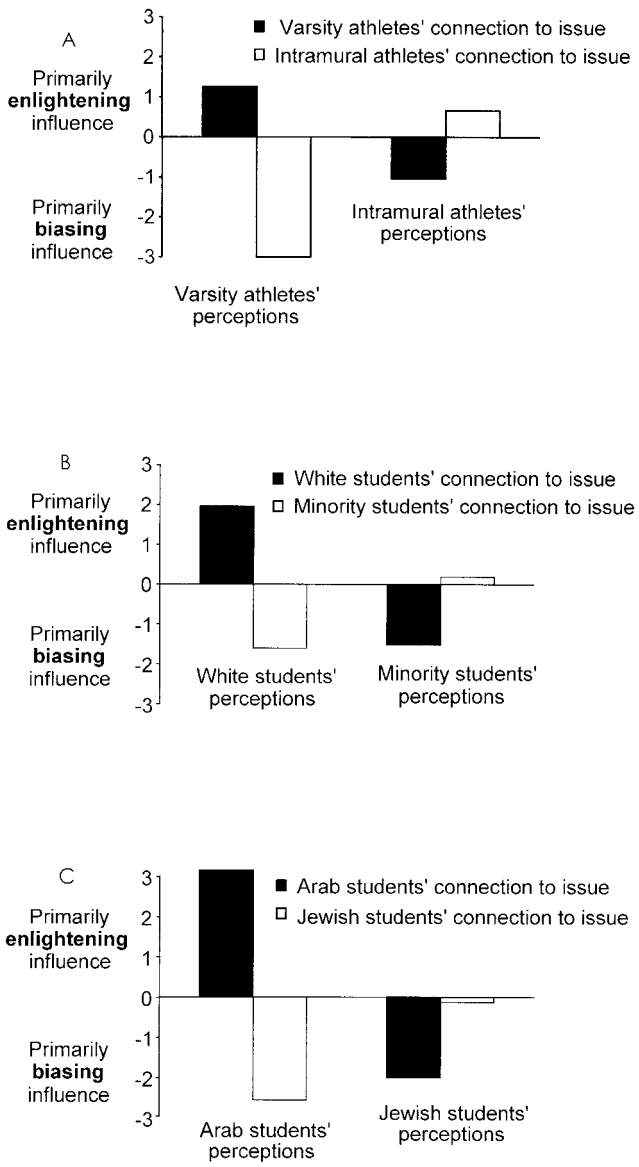

Figure 4. Perceptions of how much group membership serves to bias (vs. enlighten) opinions of members of one's own group and of members of a group with an "opposing" point of view, including (a) varsity and intramural athletes' perceptions of their own and each other's opinions about a proposal to open up the varsity-only weight room to all campus athletes, (b) white and minority students' perceptions of their own and each other's opinions about the value of affirmative action, and (c) Arab and Jewish students' perceptions of their own and each other's opinions about the ongoing "Middle East crisis" (from Ehrlinger et al., 2004).

To do otherwise would be perverse, although there clearly are cases in which individuals are willing to acknowledge that their decisions were biased, or at least influenced, by considerations that some peers would see as nonnormative or even irrational (e.g., sentimentality, or respect for the wishes of some constituency). Nevertheless, as we try to make clear in the General Discussion, the belief that one's own perceptions, beliefs, priorities, and policy preferences are the most reasonable ones in light of the information at hand has consequences that are nonobvious and important.

The Role of Introspection in Perceptions of Own Versus Others' Bias

When induced to consider the possibility that one has been less than objective and open minded, it is natural for one to introspect about the matter. But, as natural and appropriate as looking inward to answer those questions might be, it is unlikely to bear fruit because the processes that give rise to most biases are unlikely to leave a phenomenological trace accessible to simple introspection. In fact, introspections often produce the honest conviction that one acted as one did in spite of, not because of, one's private sentiments. Van Boven, White, Kamada, and Gilovich (2003) reported research evidence supporting this contention. Participants in their studies claimed to be less susceptible than others to the correspondence bias in part because they thought they engaged in more energetic efforts than others to "correct" or moderate their initial dispositional inferences by taking the prevailing situational constraints into account.

Further evidence linking blindness to one's biases to the illusion that introspection should provide evidence of such biases comes from a study in which the assessments of participants who have just manifested a particular bias—and thus enjoyed an opportunity both to note their own response and to reflect on the basis for that response—are compared with the assessments of participants who have been asked abstractly about their susceptibility to that bias without their having made any recent response that manifested it (Ehrlinger et al., 2004). One group of participants first rated themselves relative to their peers on a series of traits expected to elicit the better-than-average effect, while a second group simply considered how they *would* rate themselves if asked to do so. All participants were then asked to indicate how much their ratings had been (or would be) driven by "a desire to be objective" and how much their ratings had been (or would be) driven by "a desire to present themselves positively."

The results indicated that participants rated themselves as *less* susceptible to the relevant self-enhancement bias when they had just assessed themselves on the various traits than when they had merely contemplated making such assessments. Furthermore, this result held when the judgments of only those participants who actually exhibited the self-enhancement effect in their earlier ratings were compared with the assessments of participants who earlier had contemplated their standing among their peers in the abstract. In other words, the experience of actually committing the bias—and considering their subjective experiences while doing so—made individuals *less* open to the hypothesis that they were susceptible to that bias.

Another line of relevant evidence comes from a recent study by Pronin, Ross, and Gilovich (2004) in which participants were asked to assess either their own susceptibility or that of their fellow students (i.e., the "average Harvard student") to one of several biases such as the fundamental attribution error or the halo effect. Each bias was described in a couple of sentences. Participants who had been asked to provide self-assessments were then asked to indicate the extent to which their assessments involved (a) trying to get "inside (their own) head" and considering whether they could find evidence of the motives or considerations that had been described and (b) trying to consider "how well the description fits (their theories about) the way that people in general tend to behave." In the case of participants who had been asked about the

susceptibility of their average peer, the two follow-up questions similarly asked them to indicate the extent to which their susceptibility assessments had been based on an attempt to "get inside the heads" of their collegiate peers and the extent to which their assessments had been based on consideration of how "people in general" behave.

The results of this direct inquiry were revealing (see Figure 5).

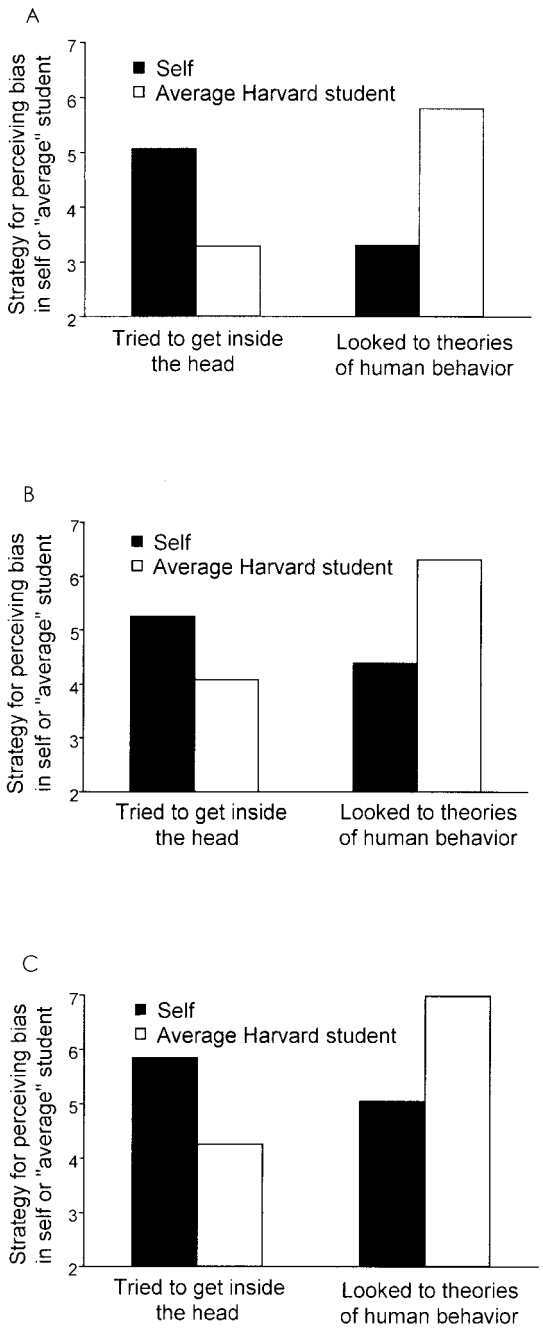

Figure 5. Self-reported use of strategies for assessing bias in oneself versus one's peers, in the case of the self-serving bias (A), the fundamental attribution error (B), and the positive halo effect (C; from Pronin et al., 2004).

When it came to assessing their own susceptibility to the relevant bias, participants reported "trying to get inside their head" to find evidence of bias far more than they reported "looking to theories of human behavior." By contrast, when it came to assessing bias in their collegiate peers, the pattern of assessments was reversed. In addition, correlational analyses revealed that greater consideration of introspective information was associated with lower perceptions of susceptibility to bias. That is, the more participants in either condition reported looking inside the head rather than looking to general theories to find evidence of a given bias, the less they concluded that the target of their judgment (i.e., self or fellow students) was in fact susceptible to that bias.

Why do individuals tend to consult general theories when attempting to assess bias in others but introspective evidence when attempting to assess bias in themselves? And to what extent does this tendency reflect a simple difference in access to the relevant introspective evidence versus a belief about the relative probative value of own versus others' introspections? We have contended that individuals afford a kind of sovereign status to their own introspections that they do not afford to the introspections of others. A follow-up study by Pronin et al. (2004) addressed this contention. In this study, Harvard undergraduates were asked to read descriptions of one particular bias (i.e., the self-serving bias) and were then provided with descriptions of two strategies (i.e., looking to internal information vs. considering general theories of human behavior) for assessing susceptibility to that bias. They then were asked how useful each strategy would be (a) when they had to assess their own susceptibility to the bias and (b) when an "average Harvard student" had to assess his or her susceptibility to that same bias.

The results of the study revealed that participants thought their own introspections had greater probative value than those of their peers. That is, participants claimed that in making self-assessments it would be more valuable for them to consult internal information than general theories, but that if others were making such self-assessments the reverse would be true.

Additional evidence relevant to our present contentions was obtained by Pronin et al. (2002). In one survey study, the investigators simply examined the association between the magnitude of the relevant asymmetry in bias perception for each bias investigated and the degree to which introspection would be likely to yield conscious traces or clues relevant to that bias. As anticipated, little if any asymmetry was observed in the case of the biases that raters assessed as likely to be accessible to introspection (either because the bias was seen as resulting from conscious decisions, e.g., strategic downward social comparisons, or because it was seen as producing consequences that make its influence felt after the fact, e.g., the tendency to underestimate the time required to complete tasks). By contrast, a marked asymmetry was apparent in the case of biases that raters deemed unlikely to be accessible to introspection (e.g., dissonance reduction, or the halo effect).

In considering the importance of introspection in assessments of bias (and also in considering motivational explanations of bias blindness), an additional finding from the Pronin et al. (2002) studies is worth noting. Participants did not see themselves as less prone than their peers to *all* negative qualities. On the contrary, they reported being quite susceptible (perhaps even more so than their peers) to negative qualities such as the tendency to procrastinate or to be fearful of public speaking—qualities, not coinci-

dentally, that are characterized by the actor's acute awareness of the relevant internal feelings. In short, although self-enhancement motives might tempt people to claim greater objectivity and freedom from bias than their peers, those motives did not seem to have an impact when the potential shortcomings were ones that were highly accessible to introspection.

General Discussion

A long progression of theory and research in psychology has advanced our understanding of the relationship between self-perception and social perception and of the ways in which they are similar and the ways in which they differ. The present attempt to understand people's tendency to see bias more readily in others than in themselves is offered as another step in that progression. This asymmetry in attributions of bias, we have suggested, arises in part from the simple fact that people inhabit a world in which others hold opinions, make judgments, and undertake decisions that differ from their own. The attempt to *account* for this difference, and to do so while holding the conviction that one's own responses to the world reflect the realities of that world, we have further suggested, is a proximate cause of the perceived asymmetry between self and others.

Our analysis further suggests that blindness to bias in the self is also produced and maintained by people's willingness to take their introspections about the sources of their judgments and decisions at face value—that is, to treat the lack of introspective awareness of having been biased as evidence that one is innocent of such bias. By contrast, when it comes to the introspections about bias offered by their peers, people are apt to doubt the authenticity and insight of those self-reports. Instead, they rely heavily on their general theories about bias and the "circumstantial evidence" at hand regarding their peers' responses and the specific incentives, constraints, motives, and beliefs that might be influencing those responses.

Do naive realism and reliance on introspective evidence ensure that individuals will never acknowledge departures from objectivity? Everyday experience suggests otherwise. Indeed, it is not uncommon to hear someone preface a remark by saying, "I may be biased, but . . ." (e.g., "I may be biased, but anyone could see that my grandson's performance in the school play showed greater verve and sophistication than the performances of his costars"). To be sure, statements like this are often made purely for public consumption. The speaker realizes that the remark is likely to be viewed with suspicion because of the circumstances and motives that attend it and wishes to acknowledge this awareness in a way that suggests the judgment is not, in fact, biased.

Beyond such disclaimers, however, there are times when one genuinely entertains the possibility that one has been biased. Consideration of the twin mechanisms responsible for the bias blind spot provides insight as to when one is likely to do so. With respect to naive realism, it is axiomatic that the question of whether one has been guilty of bias is most pressing when one's opinion differs not from a single other person but from a group of people, particularly if they are recognized to be better informed than oneself and if they represent a diversity of views rather than an identifiable bloc. With respect to introspection, some biases are likely to leave more of a phenomenological trace than others. Looking inward, for example, is not likely to reveal evidence of bias when one has fallen prey to cognitive biases such as availability, anchoring, or confirmatory hypothesis testing. But the tug of some motivational biases such as self-interest or rationalization can at times register phenomenologically and thus leave just the sort of trace to support the suspicion of bias.

By the same token, we all have experienced cases in which people do acknowledge, after the fact, instances in which they have been guilty of prejudice, favoritism, wishful thinking, or excessive attention to their own best interests. Indeed, in none of our studies did participants claim total imperviousness to bias. They simply claimed that they had been less susceptible than the "average" person or particular peers—especially peers who had made an inference, rendered a judgment, or made a decision that differed from their own. We also contend that if future research uncovers bias-to-bias differences in the permeability of the blind spot, the differences will be predictable from the extent to which the different biases leave phenomenological traces.

Consideration of the limits of introspection and the core tenets of naive realism suggests a number of related self–other differences. In particular, asymmetric (and invidious) self–other comparisons are apt to be seen in assessments about the impact of nonnormative influences beyond the classic perceptual, cognitive, and motivational biases we have examined in this article. We conclude by discussing a number of these related phenomena and considering the implications of our analysis for intergroup conflict and dispute resolution.

The "Third-Person Effect"

Vast amounts of money are spent on media attempts to influence viewers' behaviors, ranging from what they buy to how they vote. But most people who have seen the typical beer commercial featuring bikini-clad models, or the classic aspirin commercial featuring the obviously paid testimonial of an actor who says that he is not a physician but he "plays one on TV," are apt to believe that, in their own case, these crude attempts at persuasion have fallen on deaf ears. At the same time, most viewers suspect that "other people" do not share their powers of rational analysis and restraint (because otherwise, the advertisers who pay for such blatant influence attempts would be "wasting their money"). Studies have found that people show such convictions about the relative susceptibility of self versus others to biased news reporting (Vallone et al., 1985) and to a range of media appeals and propaganda (e.g., Duck & Mullin, 1995; Gunther & Thorson, 1992; Hoorens & Ruiter, 1996; Innes & Zeitz, 1988; Perloff, 1993).

In one study of this *third-person effect* (Innes & Zeitz, 1988), participants were stopped on the street and asked to judge the likely impact of three media presentations on themselves versus other respondents. For all three presentations—a political campaign conducted by a major political party in the region of Australia where the study was conducted, a depiction of media violence, and a campaign designed to deter people from associating with individuals who drink and drive—participants rated others as more likely to be influenced than themselves.

Other studies of the third-person effect suggest the operation of the specific mechanisms we have discussed. Duck and Mullin (1995) and Hoorens and Ruiter (1996) report that participants see others as more susceptible than themselves to influences that are deemed unreasonable or undesirable (e.g., sexist media presenta-

tions) but not to messages that are deemed rational and worthy of attention (e.g., advertisements supporting organ donation). In other words, people think they respond appropriately to the objective merits of the relevant messages and worry that some of their peers will be less "realistic." Gunther and Thorson (1992) found that participants generally see others as more susceptible than themselves to media appeals, but they do not make this claim about emotion-based appeals—that is, those appeals that tug at their heart strings and thus produce clear phenomenological traces of the potential source of influence.

Although the general conviction that we are less susceptible to media influence than our peers arises in part from specific instances in which we see our peers succumb (or believe they would succumb) to a message we are confident has had no impact on us, the analysis developed in this article suggests that this conviction also arises in part from a simple matter of definition. That is, when we find a message persuasive, we feel that we have responded appropriately or normatively (and that our peers will, or at least should, do likewise if they are sufficiently discerning); we do not feel that we personally have succumbed to a message that more discerning peers—or less gullible and vulnerable peers—would ignore.

As with the perceived asymmetry in perceptions of bias that we have discussed, introspection generally enhances the third-person effect. In particular, we know that we *intend* not to be influenced by appeals to prejudice, unfulfilled psychic needs, or simple vanity. (And previous research by Kruger & Gilovich, 2004, shows that people give themselves more credit than they give others for these sorts of "good intentions.") However, we give little credence to others who claim that their seemingly prejudiced judgment was not the product of such appeals, and we reject accusations that we ourselves have fallen prey to those appeals when our introspections yield justifications that seem innocent of prejudice.

The Introspection Illusion and Asymmetric Impressions of the "True" Self

The tendency to treat introspections as decisive data for self-understanding, while ignoring or responding skeptically to others' reports about their inner lives, has an important consequence. It leads us to see ourselves as best defined and most authentically represented by our private thoughts and feelings but to see our peers as best defined by their overt actions (e.g., Andersen, 1984; Andersen & Ross, 1984). By the same token, people are apt to show an *illusion of asymmetric insight*, or the conviction that they know their peers better than their peers know them (Pronin, Kruger, Savitsky, & Ross, 2001). This illusion is rooted in a pair of convictions involving the relative importance of introspection for self versus other. The first is that knowing *us* demands that one enjoy access to our private thoughts, feelings, motives, intentions, and so forth. The second is that we can know *others* quite well solely from paying attention to their behaviors, gestures, verbal responses, and other observable manifestations.

Naive Realism and the Broader Canvas of Everyday Biases and Illusions

Our account of the asymmetry in perceived bias is best appreciated when considered within the broader framework of naive realism. The essential component of naive realism—namely, the (false) sense that one sees the world as it is and not as it is filtered through one's expectations, needs, or motives, or "constructed" by one's sensory and cognitive apparatus—is presented in Figure 6. Also presented there are the two components of naive realism that follow immediately from this core conviction: the expectation that "objective and reasonable others" will share one's perceptions and judgments, and the inference that those who do not share one's perceptions and judgments are therefore either uninformed, biased, or under the influence of idiosyncratic dispositions. Finally, Figure 6 presents a number of more specific attributional and information-processing phenomena that stem from these three basic components of naive realism, some of which we shall enlarge upon briefly here.

When others see things differently. When confronted by an individual whose responses differ from one's own, the naive realist faces an attributional dilemma: *Why does this person see the world so differently?* Sometimes this dilemma is resolved in the fashion emphasized by Jones and Nisbett (1972): by concluding that the other person possesses some telling disposition (or set of dispositions) that is dictating his or her reactions. This inference is particularly likely when the discrepancy involves not a differing viewpoint or interpretation but a behavior that differs from the way one would behave (or anticipates one would behave). The tendency to resolve this attributional dilemma by drawing a dispositional inference is, in turn, related to the correspondence bias (Jones, 1990) or fundamental attribution error (Ross, 1977). Often, this error results from the incorrect and insufficiently examined assumption that the person who behaves differently is in fact responding to the same situation one has in mind oneself. By assuming that the other person is responding differently to the same "object of judgment" as oneself (Asch, 1952), one fails to appreciate the true situational constraints governing the actor's behavior and thus runs the risk of drawing an unwarranted dispositional inference.

As Figure 6 makes clear, the attributional dilemma set in motion by responses discrepant from one's own is often resolved in other ways. Most charitably, perhaps, people may assume that those others are misinformed or have not had sufficient contact with the pertinent information. This motivates attempts to "set the other person straight," accompanied by the expectation that such efforts at persuasion will prove easy and should proceed smoothly to a successful conclusion. When that expectation is not borne out, less benign attributions result, including the attribution of bias that we have emphasized throughout this article.

We have focused thus far on the tendency to infer bias on the part of specific individuals or groups—especially one's ideological adversaries—who are apt to be *defined* (in one's mind, at least) by the ways in which they fail to share one's perceptions, assessments, and judgments. But as Figure 6 illustrates, the same attribution of bias extends to mediators, news reporters, and other third parties whose job it is to offer neutral and unbiased views. The basis for such attributions of media hostility and bias is straightforward. If one's own views are experienced as ineluctable products of objective perception of the relevant issues or events (especially if one sees the world in "black or white"), such third parties (to the extent they claim the world to be a shade of gray) will be seen as biased in favor of the "other side" (Vallone et al., 1985).

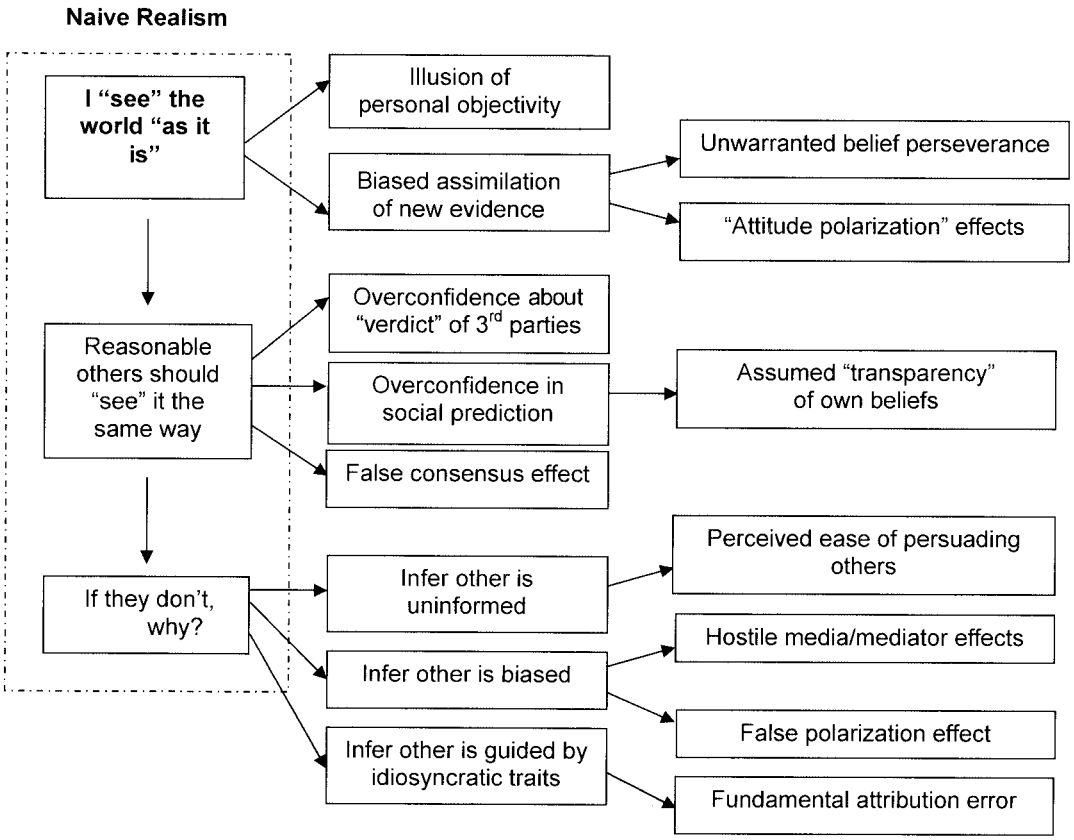

Figure 6. Convictions of naive realism and resulting consequences and phenomena relevant to conflict and misunderstanding.

Assuming that others will see the world as we see it. Additional inferential shortcomings that have been frequently documented in the social judgment literature result not from the attributional dilemma that occurs when one realizes that someone else sees the world differently from oneself, but from the prior and more basic assumption that any reasonable or unbiased person will see it the same as oneself. Most obvious is the presumption that disinterested third parties (including judges or juries) will share one's assessments about the relative merits of the two sides in a case (Babcock & Loewenstein, 1997). Another such phenomenon—the so-called *false consensus effect* (Ross, Greene, & House, 1977)—involves a tendency for personal preferences or political positions to be deemed more common by those who hold them than by those who hold alternative views. Thus, for example, people who prefer an Australian shiraz to a California merlot are likely to estimate that a higher percentage of wine lovers would likewise prefer the shiraz to the merlot than would individuals with the opposite preference.

Ross et al. (1977) attributed the false consensus effect to the same failure to distinguish between subjective perceptions and objective reality that we have emphasized—namely, the tendency to assume that one's own take on the world is a faithful reflection of objective reality and thus will be shared by others. Gilovich (1990) provided evidence supporting this "construal interpretation," finding larger false consensus effects on items that permit wide latitude for personal interpretation of the relevant alternatives (e.g., whether one has difficulty controlling one's temper) than on items that permit little latitude for divergent interpretation (e.g., whether one prefers to watch football or basketball games). Similarly, larger false consensus effects tend to be observed when choices are presented in a relatively abstract way (e.g., "Do you prefer the color tan or aqua?") than when they are presented more concretely (e.g., "Do you prefer this particular tan color swatch or this aqua one?"). These findings make clear that the false consensus effect is due at least partly to individuals' failure to recognize and make allowance for the fact that their own preferences follow from their personal construals of the matters at hand, and others might very well arrive at different construals and thus express different preferences.

This failure to anticipate that others may construe a question, issue, opportunity, or context differently than oneself also plays a significant role in a second much-researched inferential shortcoming: the tendency to make overly confident predictions about behavior, both about others' behavioral choices (Dunning, Griffin, Milojkovic, & Ross, 1990) and about one's own (Vallone, Griffin, Lin, & Ross, 1990). When predicting whether another individual is likely to enjoy an office holiday party, to prefer comedies to dramas, or to attend a peace rally, people try to take into account their knowledge (or make allowance for their lack of knowledge) about that individual's particular traits, tastes, or opinions. What

they do not make sufficient allowance for is the possibility that these "objects" of judgment (Asch, 1952) will be construed differently by the relevant actor—that the office party may be construed not as a chance to socialize with one's peers but as an opportunity to ingratiate oneself with the boss, or that the peace rally may be construed not as an effort to influence policy but as an attempt to show one's pacifist grandparent that her attempts at influence have not totally been in vain.

Of course, what applies to predictions of another person's behavior applies as well to predictions we make about our own behavior. We must decide whether to spend a sabbatical in Paris, Sydney, or Doha by constructing a representation of what each city is like. But the France, Australia, or Qatar we actually experience may differ markedly from the one we imagined, and forecasts that do not make adequate allowance for such mistaken construals are forecasts that are likely to be held with more confidence than will prove warranted as events unfold.

Assuming that one sees the world the way it is. As the uppermost boxes in Figure 6 convey, there are also biases and inferential shortcomings that stem directly from the conviction that one sees the world in a direct, "unmediated" fashion. The most obvious of these involves the subjective impression that one's own reactions and views are the product of objective analysis rather than motivational or cognitive bias—an impression that helps to produce and sustain the asymmetry in perceptions of bias in others versus self that has been this article's focus. It is in this impression that naive realism and the introspection illusion come together. The experience we have in perceiving and evaluating the world around us is that of a direct and unmediated process, and the failure of introspection to uncover any phenomenological trace of mediation reinforces the sense that we see things "as they are" and that we make judgments accordingly.

The sense that our understanding of situations and events is veridical has implications for the way we evaluate evidence pertinent to our beliefs. Evidence consistent with what we assume to be true is accepted with little scrutiny, while evidence that contradicts our beliefs thereby challenges our understanding and is subjected to intense scrutiny to resolve the epistemological discrepancy. Often this results in discounting the relevance of the contradictory information altogether or at least minimizing its import (Ditto & Lopez, 1992; Edwards & Smith, 1996; Lord et al., 1979). This biased assimilation of new information, in turn, leads to unwarranted perseverance of beliefs (see Ross & Lepper, 1980) and even, on occasion, to the normatively indefensible tendency for individuals on opposite sides of an issue to become more convinced of the validity of their position after being exposed to the same (mixed) body of evidence (Lord, Ross, & Lepper, 1979).

Implications for Conflict, Misunderstanding, and Dispute Resolution

Differences of opinion may be inevitable, but such differences do not inevitably and directly lead to conflict, mistrust, and pessimism about finding common ground. Often, it is the attributions placed on those differences of opinion that create or exacerbate such difficulties (see Ross & Ward, 1996). The processes we have described that lead people to feel that they are just about as liberal in their politics or just about as avant-garde in their tastes or just about as diligent in performing their duties as it is "reasonable" to be (i.e., that those who are more liberal are guilty of "wishful thinking," that those who are more avant-garde are guilty of "affectation," and that those who are more diligent are guilty of "false consciousness") may do little harm. The conviction that it is the other party alone who is guilty of bias in the context of a given disagreement between spouses, neighbors, or coworkers may lead to some ill feelings—especially if the parties express that sentiment aloud—but those feelings are apt to subside as shared interests and the benefits of mending fences come to the fore. But in the intergroup conflicts around the world today, the conviction that one's own side has a monopoly on objectivity, and that it is only the other side that refuses to see the past or the present as it really is, can have serious consequences indeed.

The convictions of naive realism can make parties feel that the other side is irrational or too biased to be reasoned with (rather than merely subject to the same cognitive and motivational biases that afflict all human beings—including oneself and one's ideological and political allies). Moreover, when the parties do air their grievances, they may conclude that the other side is being "strategic" and doesn't really believe what it is saying. Or perhaps worse, the conclusion reached may be that the other side really does believe what it is saying and that a rational, interest-based negotiation will thus prove fruitless, and that therefore only the use of force can win one's own side the just outcome to which it is entitled. Of course, when force is applied, the cycle of aggression and retaliation is apt to be self-perpetuating, as both sides see the other as the instigator and their own side as acting in self-defense or trying to teach the other side a lesson.

Naive realism also plays a role in creating what Robinson et al. (1995; also Robinson & Keltner, 1996) have labeled *false polarization*—that is, the perception of a greater gap between antagonists than is actually the case. This phenomenon (see Figure 6) arises from the inference that the other side's views reflect top-down bias rather than bottom-up reasoning from the "true facts," and therefore that the other side's views will be more one-sided and less nuanced than one's own. This research focused on domestic disputes within the United States about matters such as abortion rights and the appropriate canon of "Great Books" for first-year college students. But it is easy to see the relevance of these findings for intergroup conflict in Northern Ireland, the Middle East, Kashmir, and many other places currently enduring violence and suffering—where undue pessimism about bridging the gap between antagonists and the possibility of finding common ground may be as much a part of the problem as the gap itself.

Naive Realism, Biased Perceptions, and Perceptions of Bias in Troubled Times

Many critics have argued that the biases so frequently documented in the psychological literature over the past 30 years are not worthy of the attention they have received—that such "biases" are the product of methodological artifact, tricky experimenters, or assessments of human capacities removed from appropriate ecological niches (Cosmides & Tooby, 1996; Gigerenzer, 1996; Krueger & Funder, in press). Because the reality and significance of the biases themselves have been questioned, we do not doubt

that the reality and significance of a self–other asymmetry in perception of such biases will be questioned as well.

However, as we reflect on a contemporary world in which ethnic and religious strife so threaten our collective futures, and accusations of irrationality, ideological bias, and naked pursuit of self-interest accompany such strife, we find the objections of such critics to be wanting. The evidence is all too plain that human inferential shortcomings—including overconfident prediction and biased assimilation of the lessons of history, compounded by people's unwillingness to consider the possibility that they are just as susceptible to those biases as those they revile—are continually and ubiquitously making their influence felt. At the very least, these shortcomings in judgment and insight serve to exacerbate and perpetuate the historical and economic roots of the conflicts that are all-too-present elements of the human condition.

References

Andersen, S. M. (1984). Self-knowledge and social inference: II. The diagnosticity of cognitive/affective and behavioral data. *Journal of Personality and Social Psychology, 46,* 294–307.

Andersen, S. M., & Ross, L. (1984). Self-knowledge and social inference: I. The impact of cognitive/affective and behavioral data. *Journal of Personality and Social Psychology, 46,* 280–293.

Armor, D. A. (1999). The illusion of objectivity: A bias in the perception of freedom from bias. *Dissertation Abstracts International, 59,* 5163B.

Asch, S. E. (1952). *Social psychology.* New York: Prentice-Hall.

Babcock, L., & Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. *Journal of Economic Perspectives, 11,* 109–126.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54,* 462–479.

Baumeister, R. F. (1998). The self. In D. T. Gilbert & S. T. Fiske (Eds.), *The handbook of social psychology* (4th ed., pp. 680–740). New York: McGraw-Hill.

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.

Bruner, J. S. (1957). Going beyond the information given. In H. Gruber, K. Hammond, & R. Jesser (Eds.), *Contemporary approaches to cognition* (pp. 41–69). Cambridge, MA: Harvard University Press.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology, 67,* 366–381.

Cooley, C. H. (1902). *Human nature and the social order.* New York: Scribner.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment and uncertainty. *Cognition, 58,* 1–73.

Dawson, E., Gilovich, T., & Regan, D. T. (2002). Motivated reasoning and the Wason selection task. *Personality and Social Psychology Bulletin, 28,* 1379–1387.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology, 63,* 568–584.

Dovidio, J. F., & Gaertner, S. L. (1991). Changes in the expression and assessment of racial prejudice. In H. J. Knopke, R. J. Norrell, & R. W. Rogers (Eds.), *Opening doors: Perspectives on race relations in contemporary America* (pp. 119–148). Tuscaloosa: University of Alabama Press.

Duck, J. M., & Mullin, B. (1995). The perceived impact of the mass media: Reconsidering the third person effect. *European Journal of Social Psychology, 25,* 77–93.

Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology, 58,* 568–581.

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 Psychology, 57,* 1082–1090.

Edwards, K., & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology, 71,* 5–24.

Ehrlinger, J., Gilovich, T., & Ross, L. (2004). *Peering into the bias blind spot: People's assessments of bias in themselves and others.* Manuscript submitted for publication.

Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments produced by errors in self- or social prediction? *Journal of Personality and Social Psychology, 79,* 861–875.

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

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 288–299.

Friedrich, J. (1996). On seeing oneself as less self-serving than others: The ultimate self-serving bias? *Teaching of Psychology, 23,* 107–109.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review, 103,* 592–596.

Gilbert, D. T., Brown, R. P., Pinel, E. C., & Wilson, T. D. (2000). The illusion of external agency. *Journal of Personality and Social Psychology, 79,* 690–700.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin, 117,* 21–38.

Gilovich, T. (1990). Differential construal and the false consensus effect. *Journal of Personality and Social Psychology, 59,* 623–634.

Gilovich, T., & Griffin, D. (2002). Heuristics and biases: Then and now. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 1–18). New York: Cambridge University Press.

Goethals, G. R. (1986). Fabricating and ignoring social reality: Self-serving estimates of consensus. In J. M. Olson, C. P. Herman, & M. P. Zanna (Eds.), *Relative deprivation and social comparison: The Ontario Symposium* (Vol. 4, pp. 135–157). Hillsdale, NJ: Erlbaum.

Gunther, A. C., & Thorson, E. (1992). Perceived persuasive effects of product commercials and public service announcements: Third-person effects in new domains. *Communication Research, 19,* 574–596.

Heath, C. (1999). On the social psychology of agency relationships: Lay theories of motivation overemphasize extrinsic incentives. *Organizational Behavior and Human Decision Processes, 78,* 25–62.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Hoorens, V., & Ruiter, S. (1996). The optimal impact phenomenon: Beyond the third person effect. *European Journal of Social Psychology, 26,* 599–610.

Ichheiser, G. (1949). Misunderstandings in human relations: A study in false social perception. *American Journal of Sociology, 55*(Suppl.).

Innes, J. M., & Zeitz, H. (1988). The public's view of the impact of the mass media: A test of the "third person" effect. *European Journal of Social Psychology, 18,* 457–463.

Jones, E. E. (1990). *Interpersonal perception.* New York: Freeman.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 220–266). New York: Academic Press.

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the cause of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution:*

Perceiving the causes of behavior (pp. 79–94). Morristown, NJ: General Learning Press.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192–238). Lincoln: University of Nebraska Press.

Krueger, J. (1998). Enhancement bias in description of self and others. *Personality and Social Psychology Bulletin, 24,* 505–516.

Krueger, J., & Funder, D. (in press). Towards a balanced social psychology: Causes, consequences and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences.*

Krueger, J., & Zeiger, J. S. (1993). Social categorization and the truly false consensus effect. *Journal of Personality and Social Psychology, 65,* 670–680.

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.

Kruger, J., & Gilovich, T. (2004). Actions, intentions, and self assessment: The road to self-enhancement is paved with good intentions. *Personality and Social Psychology Bulletin, 30,* 328–339.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology, 53,* 636–647.

Lieberman, M. D., Ochsner, K. N., Gilbert, D. T., & Schacter, D. L. (2001). Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science, 121,* 135–140.

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.

Miller, A. G., Baer, R., & Schonberg, P. (1979). The bias phenomenon in attitude attribution: Actor and observer perspectives. *Journal of Personality and Social Psychology, 37,* 1421–1431.

Miller, D. T., & McFarland, C. (1987). Pluralistic ignorance: When similarity is interpreted as dissimilarity. *Journal of Personality and Social Psychology, 53,* 298–305.

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.

Nisbett, R. E., Caputo, C., Legant, P., & Marecek, J. (1973). Behavior as seen by the actor and as seen by the observer. *Journal of Personality and Social Psychology, 27,* 154–164.

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

Nisbett, R. E., & Wilson, T. D. (1977a). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 8,* 231–259.

Nisbett, R. E., & Wilson, T. D. (1977b). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35,* 250–256.

Perloff, R. M. (1993). Third-person effect research 1983–1992: A review and synthesis. *International Journal of Public Opinion Research, 5,* 167–184.

Pronin, E., Kruger, J., Savitsky, K., & Ross, L. (2001). You don't know me, but I know you: The illusion of asymmetric insight. *Journal of Personality and Social Psychology, 81,* 639–656.

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin, 28,* 369–381.

Pronin, E., Ross, L., & Gilovich, T. (2004). *Strategies for assessing bias in self versus others*. Manuscript in preparation.

Pyszczynski, T. A., & Greenberg, J. (1981). Role of disconfirmed expectancies in the instigation of attributional processing. *Journal of Personality and Social Psychology, 40,* 31–38.

Robinson, R. J., & Keltner, D. (1996). Much ado about nothing? Revisionists and traditionalists choose an introductory English syllabus. *Psychological Science, 7,* 18–24.

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. (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). New York: Academic Press.

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. (1980). The perseverance of beliefs: Empirical and normative considerations. In R. A. Shweder & D. Fiske (Eds.), *New directions for methodology of behavioral science: Fallible judgment in behavioral research* (pp. 17–36). San Francisco: Jossey-Bass.

Ross, L., McGuire, J., & Minson, J. (2004). *The relationship between self-other disagreement and the perceived impact of biasing versus normative considerations on own versus others' opinions*. Manuscript in preparation.

Ross, L., Pronin, E., & Lin, D. Y. (2004). *The bias blind spot in intergroup perception.* Manuscript in preparation.

Ross, L., & Ward, A. (1995). Psychological barriers to dispute resolution. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 27, pp. 255–304). San Diego, CA: Academic Press.

Ross, L., & Ward, A. (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. In T. Brown, E. S. Reed, & E. Turiel (Eds.), *Values and knowledge. The Jean Piaget Symposium Series* (pp. 103–135). Hillsdale, NJ: Erlbaum.

Savitsky, K., & Gilovich, T. (2003). The illusion of transparency and the alleviation of speech anxiety. *Journal of Experimental Social Psychology, 39,* 618–625.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261–302). New York: Academic Press.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103,* 193–210.

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

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90,* 293–315.

Vallone, R. P., Griffin, D. W., Lin, S., & Ross, L. (1990). Overconfident prediction of future actions and outcomes by self and others. *Journal of Personality and Social Psychology, 58,* 582–592.

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., Dunning, D., & Loewenstein, G. (2000). Egocentric empathy gaps between owners and buyers: Misperceptions of the endowment effect. *Journal of Personality and Social Psychology, 79,* 66–76.

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.

Van Boven, L., White, K., Kamada, A., & Gilovich, T. (2003). Intuitions about situational correction in self and others. *Journal of Personality and Social Psychology, 85,* 249–258.

Vivian, J. E., & Berkowitz, N. H. (1992). Anticipated bias from an outgroup: An attributional analysis. *European Journal of Social Psychology, 22,* 415–424.

Wetzel, C. G., Wilson, T. D., & Kort, J. (1981). The halo effect revisited: Forewarned is not forearmed. *Journal of Experimental Social Psychology, 17,* 427–439.

Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious.* Cambridge, MA: Belknap/Harvard.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin, 116,* 117–142.

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General, 125,* 387–402.

Wong, P. T. P., & Weiner, B. (1981). When people ask "why" questions, and the heuristics of attributional search. *Journal of Personality and Social Psychology, 40,* 650–663.

Received October 29, 2002
Revision received August 1, 2003
Accepted August 13, 2003 ■