

AMIRAM RAVIV (345), Department of Psychology, Tel Aviv University, Tel Aviv 69978, ISRAEL

DAVID O. SEARS (95), Psychology Department, University of California Los Angeles, Los Angeles, CA 90095

KEREN SHARVIT (345), Department of Psychology, Tel Aviv University, Tel Aviv 69978, ISRAEL

TOM R. TYLER (151), Department of Psychology, New York University, New York, NY 10003

NOTICE

This material may be protected by
Copyright Law (Title 17 U.S. code)

ACCURACY IN SOCIAL PERCEPTION: CRITICISMS, CONTROVERSIES, CRITERIA, COMPONENTS, AND COGNITIVE PROCESSES

Lee Jussim

This article reviews controversies that have surrounded research on interpersonal accuracy for the last 50 years by providing and justifying a probabilistic conception of accuracy and by addressing the following issues: Whether accuracy can be distinguished from logical coherence; whether accuracy research provides insight into cognitive processes; why accuracy in perception of a behavior or trait must be distinguished from accuracy in explanations for that behavior or trait; how to distinguish self-fulfilling prophecy from accuracy; why distinguishing among several levels of analysis is crucial with respect to evaluating the accuracy of social stereotypes; whether accuracy research helps create or alleviate social problems; how accuracy can sometimes lead to discrimination; similarities and differences between establishing scientific construct validity and the accuracy of lay social beliefs; and the nature of some of the contributions to understanding social perception provided by componential and noncomponential approaches to the study of accuracy.

What could be a more basic or obvious purpose of social perception research than assessment of the accuracy of people's perceptions of one another? And what could be simpler? Although both questions are phrased rhetorically, in fact, accuracy was an all-but-dead topic within social psychology for roughly 30 years, from about 1955 to 1985. It turned out that not only was the study of accuracy less simple than it seemed but it was, in fact, widely viewed as a theoretical and political minefield.

This chapter reviews the major controversies surrounding accuracy research and resolves those controversies, shows how researchers who have been studying accuracy have resolved those controversies, shows how identical issues have been resolved in other areas of psychology, or partially

resolves these issues. In the case of partial resolutions, this review also highlights the remaining unresolved issues and discusses how accuracy can be productively studied despite some residual difficulties.

I. What Is Social Perceptual Accuracy?

Social perceptual accuracy, in this review, refers to correspondence between perceivers' beliefs (expectations, perceptions, judgments, etc.) about one or more target people and what those target people are actually like, independent of those same perceivers' influence on them. This definition may appear straightforward, but it involves subtle issues and complexities that warrant explication—which is provided next.

A. PROBABILISTIC REALISM

This definition of accuracy is based on a set of principles referred to here as probabilistic realism. The main ideas of probabilistic realism, adopted throughout this chapter, are that there is an objective reality out there that, flawed and imperfect though we (both scientists and lay people) may be, we can eventually come to know or understand, at least much of the time. The term "probabilistic realism" is used as a somewhat simpler name for what has been called "critical realism" and "pancritical rationalism." It is essentially the same approach described by Funder (1995, 1999), which itself was heavily influenced by mainstream psychological approaches to construct validity (e.g., Cook & Campbell, 1979; Cronbach & Meehl, 1955).

Most psychological researchers implicitly adopt the idea that they can come to understand much about other people nearly all of the time, at least in their research. This perspective is often implicit in the sense that, although its assumptions are widely adopted (as shall soon be demonstrated), few research articles include sections that make explicit the authors' assumptions regarding the nature of social reality. The next section, therefore, explains both the "realism" and "probabilistic" aspects of probabilistic realism and then briefly describes what accuracy means in this context.

1. Realism

"Realism" refers to the idea that there is an objective reality that is independent of social perception. Few social scientists, except the most radical of social constructivists, deny the existence of such a reality. Whenever

empirical researchers measure any psychological variable, they implicitly assume that variable is real. Even the three-decade exile (roughly 1955–1985) of accuracy research within social psychology occurred because of heightened interest in bias and recognition of genuine complexities in studying accuracy—not because researchers argued there was no social reality. The idea that objective social reality can influence social perception and constrain the potential for bias is explicit in many theoretical perspectives within psychology (e.g., Allport, 1955; Brophy, 1983; Brunswik, 1952; Festinger, 1957; Funder, 1987, 1995, 1999; Gibson, 1979; Jussim, 1991; Kelly, 1955; Kunda, 1990; McArthur & Baron, 1983; McCauley, Stitt, & Segal, 1980). The "realism" part of probabilistic realism reflects this assumption.

2. Social Reality and Social Beliefs Are Often Inherently Probabilistic

The term "probabilistic" is used to capture three different aspects of accuracy. First, it means that most criteria are probabilistic, not definitive. This applies to perceptions both of individuals and of groups. For example, a child with a high IQ is likely to do well in school, but there is no guarantee. Similarly, if the Nepalese are more courageous than other people, this does not necessarily mean that every person from Nepal, but no one from Denmark, will rush into the palace to save the Royal Family from a rampaging prince with a gun. People from Nepal do not need to always be braver than other folks, but if the Nepalese are more courageous than other people, they will act more courageously more of the time than do others.

Second, "probabilistic" captures the idea that many social beliefs themselves are inherently probabilistic. The belief that Michael Jordan was the best basketball player of the 1990s does not require believing that his teams would win every game they played. Although people rarely phrase their beliefs in explicitly probabilistic terms, this belief is best interpreted as meaning something like "all other things being equal, having Jordan on your team will enhance your chance of winning more than having any other player on your team." Thus, many social beliefs, themselves, are probabilistic.

Stereotypic beliefs, too, are usually inherently probabilistic (see also Krueger, 1996; McCauley et al., 1980). If John believes that Asians are wealthier than other people, this does not necessarily mean that John expects all Asians to be fabulously rich or denies the existence of a single impoverished Asian. It may only mean that, on average, John believes they are richer than other people. Of course, even absolutist beliefs can be viewed as probabilities. The belief that all Englishmen are dignified can be translated into the belief that 100% of Englishmen are dignified.

Not only is social reality probabilistic, and not only are many social beliefs probabilistic, but the relationship between belief and reality is often

probabilistic. For example, if Jorge expects Rose to be late for all of their meetings, and Rose is late only 95% of the time, Jorge is still quite accurate—certainly far more accurate than had he expected Rose to generally be on time. Accuracy is rarely all or none; it is usually a matter of degree.

B. ACCURACY

Accuracy is greater the greater the correspondence between perceivers' beliefs (expectations, perceptions, judgments, etc.) about one or more target people and what those target people are actually like, independent of those same perceivers' influence on those targets. This definition has three core ideas: correspondence, what people are actually like, and independent of influence. If Bella's teacher, Mr. Jones, predicts that Bella will receive an A on her next math test, and she does, Jones's belief corresponds well with the outcome. Similarly, near misses involve closer correspondence than wildly inaccurate misses. If Jones predicted an A for Bella, and she receives a B+, Jones is still more accurate than had he predicted that she would receive a C. Accuracy as degree of correspondence between subjective perception and objective reality characterizes most modern and many older approaches to accuracy in perception (both object and social) and memory (e.g., Allport, 1958; Brunswik, 1952; Funder, 1995; Gibson, 1979; Judd & Park, 1993; Jussim, 1991; Kenny, 1994; Koriat, Goldsmith, & Pansky, 2000).

"Independent of influence" is also conceptually straightforward. It means that the perceiver cannot have caused the target's outcome. If a coach predicted that one person will become the team's best player, and the coach caused that person to become the best player (e.g., by giving that player extra time and attention, being extra supportive and encouraging to that player but not to other players, etc.), this is a self-fulfilling prophecy, not accuracy. For a perceiver's belief to be accurate, it must correspond to targets' behavior or attributes without the perceiver having caused them.

The phrase "actually like" appears simple but implies several important assumptions made explicit here. Probabilistic realism assumes that people have some characteristics that are independent of any particular perceiver's judgments of those characteristics (see Funder, 1995, 1999, 2001; Jussim, 1991; McCauley et al., 1980). Second, how to identify what those characteristics are, independent of subjective interpersonal judgment, is in essence, the criterion issue. Because identifying criteria for establishing accuracy has so often been framed as something so difficult as to threaten the viability of accuracy research altogether, it constitutes one of the major issues around which this review is focused and will be addressed in detail later.

C. WHY ACCURACY IS RARELY ABSOLUTE

1. *Is a Social Judgment That Is 99% Correct Completely Wrong?*

To consider accurate only those lay judgments that are absolutely perfect bullseyes would be to render the concept almost useless, lead to an untenable system of classifying beliefs, and reflect a scientifically embarrassing double standard. For example, a prediction that is 99% correct and one that is 0% correct would both have to be considered "inaccurate." This is useless at best and dysfunctional and misleading at worst. If, for example, one finds a money manager who can successfully predict stock moves 99% of the time, it would be most dysfunctional to dismiss the predictions as inaccurate. Indeed, one might consider emptying one's bank accounts and mortgaging the house, because this manager can make one rich very quickly.

2. *Does a Theory Need to Be Confirmed 100% of the Time to Be Considered Valid?*

Furthermore, absolute perfection constitutes a standard that scientists, including but not restricted to social psychologists, do not apply in their own work. Consider a psychological theory whose predictions were confirmed 99% of the time. By the absolutist standard for accuracy, the theory would have to be considered "inaccurate." This is clearly misleading.

Far more common is a situation like the following: Results confirm the hypotheses on two out of three dependent variables in study 1, work only among women in study 2, and then reappear on three out of four dependent variables in study 3 for both men and women. Such a study will most likely get published, and the tone will emphasize confirmation of the author's theory. Thus, this probabilistic conception of social perceptual accuracy applies a similar probabilistic standard when evaluating the accuracy of lay perceptions that scientists apply when testing their own theories and hypotheses and has been widely adopted in theory and research on accuracy in perception, social perception, and memory (e.g., Brunswik, 1952; Funder, 1995; Jussim, 1991; Kenny, 1994; Koriat et al., 2000; McCauley et al., 1980; Ryan, 2002).

II. Why Is Accuracy Research Valuable?

But why even bother? Before even attempting to resolve controversies or navigate minefields, it might be worthwhile to explicitly state why some scientists consider research on accuracy to be important.

A. SOME CONTRIBUTIONS OF ACCURACY RESEARCH

Many scientists perform accuracy research because they are interested in understanding the basic social and psychological processes underlying interpersonal interactions, close relationships, motivation and achievement, and intergroup relations (e.g., Ambady & Rosenthal, 1992; Funder, 1995; Ickes, 1997; Judd & Park, 1993; Kenny, 1994; Lee, Jussim, & McCauley, 1995). Accuracy research over the last 20 years has addressed issues such as how and why the length of relationships influences social perceptual accuracy; how people identify and integrate behavioral cues to reach judgments about others' personal characteristics; whether stereotypes typically exaggerate, underestimate, or accurately describe group differences; and whether teacher expectations predict student achievement primarily because of accuracy, bias, or self-fulfilling prophecy (and this incomplete list is provided mainly to provide some sense of the issues addressed by accuracy research, not to comprehensively describe the entire body of work—see, e.g., Archer & Akert, 1977; Dawes, 1979; Goldman & Lewis, 1977; Hall & Carter, 1999; Koriati et al., 2000 for accuracy research on other topics).

B. BALANCE?

Another reason that research on accuracy is important is to evaluate the validity of the stark, dark impression of human judgment, perception, and thinking that has emerged from the literature on social cognition and decision making. Social psychology has a long history of emphasizing the evil, immoral, and irrational aspects of human thoughts, feelings, and behaviors (see, e.g., Krueger & Funder, 2004, for a review; or the discussions that appear in almost every social psychology text regarding lack of bystander intervention in emergencies, conformity and obedience, the fundamental attribution error, self-serving biases, and judgment under uncertainty).

1. *Social Psychological Scholarship Convinces Undergraduates That People Are Fundamentally Irrational*

When I teach social psychology as a small seminar, I offer the following question as an essay on the midterm: "Does social psychological research indicate that people are fundamentally logical and rational or fundamentally illogical and irrational?" The required readings up to the midterm include Myers's (1999) introductory social psychology text's chapter on the self,

Aronson's (1999) *Social Animal* chapter on self-justification, Merton (1948) on self-fulfilling prophecies, La Piere's (1934) attitudes-do-not-predict-behavior study, and two chapters each from Cialdini's (1993) book on social influence and Ross and Nisbett's (1991) book on the person and the situation. Although not a "random sample," these are, I think, well-respected, common, and mainstream social psychological writings.

Consistently, about three quarters of my students conclude that social psychology indicates that people are fundamentally irrational. Consider the following quotations from three papers:

First student, introductory sentence: "Through taking this class, I have come to the conclusion that people are, and have always been, primarily irrational."

Second student, introductory sentence: "People are not rational beings; rather they are rationalizing beings."

Third student, concluding sentence: "I guess that we are probably irrational and spend our lives trying to convince ourselves that we are rational."

Intentionally or not, social psychological scholarship often creates the impression that people are fundamentally irrational.

2. *Some Prominent Scientists Have Reached the Same Conclusion*

This impression is not restricted to undergraduates exposed to social psychology for the first time. Consider the following:

"several decades of experimental research in social psychology have been devoted to demonstrating the depths and patterns of inaccuracy in social perception. . . . This applies . . . to most empirical work in social cognition" (Jost & Kruglanski, 2002, pp. 172–173).

"Social perception is a process dominated far more by what the judge brings to it than by what he takes in during it" (Gage & Cronbach, 1955, p. 420).

"Our beliefs pervasively color and bias our response to subsequent information, evidence, or argumentation" (Lord, Lepper, & Preston, 1984, p. 1231).

"Stereotypes are maladaptive forms of categories because their content does not correspond to what is going on in the environment" (Bargh & Chartrand, 1999, p. 467).

And, of course, several highly influential books have been written on error and bias, which, except for an occasional passing acknowledgment that people are not always wrong, primarily focus on how judgment and perception go astray (Gilovich, 1991; Kahneman, Slovic, & Tversky, 1982; Nisbett & Ross, 1980).

3. *Is the Glass Really Almost Empty? Evaluating the Validity of Perspectives Emphasizing Human Bias and Irrationality*

In this context, then, research on accuracy becomes important for either of two reasons. One reason is that many social psychologists may emphasize error and bias because they believe that social judgment really is overwhelmed by foolishness and irrationalities. Those subscribing to this view may not completely deny the possibility of accuracy or rationality, but the quotations above indicate that some highly influential social psychologists seem to conclude that inaccuracy, bias, and irrationality are the norm.

To the extent that research within this perspective is interpreted broadly as “demonstrating the depths and patterns of inaccuracy in social perception” (Jost & Kruglanski, 2002), then work on accuracy is essential for evaluating its validity. To reach broad and general conclusions about inaccuracy, one cannot simply study error and bias. If studies are designed just to test for bias, then they will either find bias (if successful) or nothing (if unsuccessful—see Krueger & Funder, 2004, for a more detailed analysis of this point). Either way, many such studies provide little or no information regarding accuracy. The only theoretically viable way to reach broad and generalizable conclusions about inaccuracy is to conduct accuracy research.

4. *Is the Glass Really Only Half-Empty? Evaluating the Validity of More Balanced Perspectives*

Another, not mutually exclusive, possibility is that some social psychological perspectives view people as often accurate and rational but, for a variety of reasons, consider error, bias, and irrationality more interesting and important. This perspective does not mean to create the impression that people are fundamentally irrational and overwhelmingly inaccurate. Proponents would probably agree that a more balanced impression of human social judgment and perception is warranted and desirable. Koriat et al. (2000, p. 522) reached a similar conclusion regarding research on the accuracy of memory:

“Second, while memory may in fact be more fallible and malleable than is assumed by the layman, it seems to us that the interest in memory illusions and false memories, spurred perhaps by real life problems, has led researchers to selectively focus on the dark side of memory, resulting in a somewhat biased picture.”

Replace “memory” with “social perception and judgment” and this quotation might also aptly describe the current status of research on social cognition. For social scientists subscribing to a more balanced view, work

focusing on accuracy is important to evaluate its validity. A balanced perspective would be largely justified if research typically found people arriving at biased and erroneous judgments to about the same extent that they reach unbiased and accurate judgments. However, such a perspective could also be unjustified in either of two opposing directions. First, perhaps, in general, people’s social beliefs, perceptions, and judgments are indeed dramatically more prone to error and bias than accuracy (consistent with the first perspective); or second, perhaps, in general, people’s beliefs, perceptions, and judgments are much more accurate than they are erroneous or biased. As Brunswik (1952) pointed out over 50 years ago, the only way to justifiably reach broad and general conclusions regarding the viability of the balanced perspective or either of the imbalanced perspectives is to conduct a great deal of research using a wide variety of methodologies that assesses the accuracy of people’s social beliefs, perceptions, expectations, and judgments in a wide variety of contexts—and only then trying to reach broad and general conclusions about the typical accuracy or inaccuracy of social judgments (see also Funder, 1987, 1995; Koriat et al., 2000 for similar points).

III. Why Addressing Controversies in Accuracy Research Is Important

The need to resolve theoretical and conceptual issues in the study of social perceptual accuracy is particularly acute. This is because, as shall be demonstrated throughout this chapter, many of the criticisms of accuracy research and even of the concept of social perceptual accuracy have typically not been framed in terms of the types of alternative perspectives and debates that have framed other areas of social psychology (e.g., do attitudes strongly or weakly predict behavior? Is self-enhancement or self-consistency the dominant self-motive? Does affect precede cognition, or does cognition precede affect? Do individuating information or stereotypes dominate person perception?). Instead, many of the criticisms of accuracy research have often been presented in such a manner as to imply or explicitly state that accuracy research is so flawed as to render the entire enterprise of dubious value, or at least not worth the effort. As a consequence, although theoretical controversies in any area often provide a rich springboard for inspiring research, in the area of accuracy, such controversies have historically tended to restrict and reduce such research by stigmatizing it.

Thus, addressing and resolving many of those controversies would seem to be a particularly valuable conceptual and theoretical undertaking. This review, however, does not merely identify those criticisms, difficulties, or

obstacles—it also critically evaluates them. As shall be seen, some (but not all) criticisms are more flawed than that which they criticize, and some (but not all) obstacles are more apparent than real. Some, however, constitute bona fide difficulties in conducting accuracy research. One purpose of this review, therefore, is to discriminate between, on one hand, the unjustified criticisms and pseudo-obstacles, and on the other, justified criticisms and genuine obstacles. When the criticisms are justified and the obstacles real, this review will discuss how they can be overcome.

To accomplish this goal, the remainder of this review centers around three major questions, each of which chronicles a set of obstacles to, difficulties in, or criticisms regarding accuracy research.

First, what are common theoretical, conceptual, and political criticisms of accuracy research? How valid are these criticisms? When valid, do they threaten the viability of accuracy research and, if so, how can such threats be addressed?

Second, is it possible to establish reasonable criteria against which to evaluate the accuracy of social perception? One of the most common criticisms of accuracy research is the claim that it is difficult to find appropriate criteria with which to assess the validity of social perception. In this review, this claim, which has often been presented as a semirhetorical indictment of accuracy research, is transformed into a scientific question: What criteria can be used to assess the validity of social perceptions and what are their limitations?

Third, must all research on accuracy perform a complex analysis of the components of social judgment? Another common criticism of accuracy research is that measures of social perceptions are heavily polluted by artifacts and biases. One set of proposed solutions to this problem involves difficult and complex statistical procedures to remove or account for those artifacts and biases. Unfortunately, however, these solutions have a long history of discouraging any accuracy research at all. Thus, this review summarizes existing componential approaches and addresses whether accuracy can be meaningfully assessed without using such procedures.

These questions are substantively different from one another: some are conceptual or meta-theoretical, others are methodological, and the last set is mathematical/statistical. Nonetheless, they are all similar in that, when they have been addressed in many prior reviews, the answers to them have frequently suggested that the obstacles to accuracy research are so formidable as to render such research of dubious value or not worth the effort. Thus, one contribution of this review is to critically evaluate the criticisms and objections to accuracy research that have been raised over the last 50 years.

In the process of addressing these questions, several substantive theoretical issues are also discussed: Whether accuracy can be distinguished from

logical coherence, whether accuracy research provides insight into cognitive processes, why accuracy in perception of a behavior or trait must be distinguished from accuracy in explanations for that behavior or trait, how to distinguish self-fulfilling prophecy from accuracy, why distinguishing among several levels of analysis is crucial with respect to evaluating the accuracy of social stereotypes, whether accuracy research helps create or alleviate social problems, how accuracy can sometimes lead to discrimination, similarities and differences between establishing scientific construct validity and the accuracy of lay social beliefs, and the nature of some of the contributions to understanding social perception provided by componential and non-componential approaches to the study of accuracy.

IV. What This Review Does Not Address

A. CORRESPONDENCE VERSUS COHERENCE

Accuracy, in this review, involves correspondence between belief and reality, not logical coherence of belief. Logical coherence refers to the internal logical consistency of people's beliefs and perceptions and is best represented by the vast amount of research on logical flaws and errors in decision making and social judgment (e.g., Gilovich, 1991; Hastie & Dawes, 2001; Kahneman et al., 1982; Nisbett & Ross, 1980). Consider someone who estimates that there is a higher chance of 1000 people dying in California because an earthquake leads to a huge flood than there is a chance of 1000 people dying in the United States because of a flood (Tversky & Kahneman, 1983). These predictions are logically incoherent, because floods in California are a mere subset of all floods in the United States, and floods caused by earthquakes are a subset of all floods. Therefore, it is impossible for a flood caused by an earthquake that kills 1000 people in California to be more probable than a flood that kills 1000 people somewhere in the United States.

(In)coherence is, at best, however, only weakly related to accuracy defined as correspondence [Wright & Drinkwater (1997) present a detailed exposition regarding an empirical example of the weak relationship between coherence and accuracy]. Let's say that the real probabilities of an earthquake causing a flood in California that kills 1000 people sometime next year is 1/10,000 and the probability of a flood killing 1000 people somewhere in the United States next year is 1/1000. Now let's consider two perceivers generating intuitive probabilities: Social Judge CBI (Coherent but Inaccurate) and Social Judge IBMA (Incoherent but Mostly Accurate). Social Judge CBI

believes the real probabilities are 70% and 90%. CBI is logically coherent, but his estimates are wildly inaccurate.

Social Judge IBMA, in contrast, believes that the probabilities are 1/1000 (California) and 1/10,000 (U.S.), respectively. IBMA's beliefs are logically incoherent, and her incoherence guarantees that her beliefs will not be perfectly accurate, because at least one of two, and perhaps both, incoherent beliefs must be inaccurate. Nonetheless, from the standpoint of accuracy as defined here (correspondence, independent of influence), IBMA's predictions are far closer to the truth—that is, far more accurate—than are CBI's.

Logical flaws, errors, and biases in the processes of judgment and decision making are extremely important, and nothing in this review is intended to, or should be interpreted as, challenging, undermining, or underestimating them. Such flaws in logic or process, however, do not usually provide much direct information about accuracy, as defined here. One can be coherent and accurate, coherent and inaccurate, incoherent and (largely) accurate, and incoherent and inaccurate. This review focuses exclusively on accuracy and does not address coherence further.

B. MAJOR EMPIRICAL ACCOMPLISHMENTS OF MODERN WORK ON ACCURACY

This chapter does not review the major empirical accomplishments of accuracy research. Reviews of, for example, accuracy in personality perception (Funder, 1987, 1995, 1999), accuracy in dyadic interactions (Kenny, 1994; Kenny & Albright, 1987), empathic accuracy (Ickes, 1993, 1997), accuracy based on minimal information (Ambady & Rosenthal, 1992), memory (Koriat et al., 2000), and accuracy in social stereotypes (Eagly & Diekmann, 1997; Judd & Park, 1993; Lee et al., 1995; McCauley et al., 1980; Ryan, 1995, 2002) can all be found elsewhere. The purpose of this chapter is to describe how the myriad of conceptual and methodological criticisms and problems with, and obstacles to, conducting accuracy research can be resolved or have already been resolved by the existing scholarship regarding accuracy.

C. THE PHENOMENOLOGY OF ACCURACY

This chapter is not about the phenomenology of accuracy. Whether lay people do or do not see accuracy in the same manner as described here is an interesting issue but is beyond the scope of this review. The definition of

accuracy here—probabilistic correspondence between belief and reality, independent of causal influence—is a scientific definition, not a description of or hypothesis about how lay people construe accuracy.

D. MORALITY AND LEGALITY

Universities cannot reject the application of a particular African American simply because one believes that, on average, African Americans score about one standard deviation lower than do Whites on their Standard Aptitude Tests (SATs). Such a rejection would clearly be immoral and illegal because the average characteristics of a group do not dictate the specific characteristics of any individual member. Nonetheless, the belief that African Americans, on average, score about 100 points lower on each of their SATs is accurate (e.g., Neisser et al., 1996). The claim about this particular applicant's qualification, however, may be—but is not necessarily—wrong. Whether the claim that this particular applicant has low SAT scores is accurate depends on the applicant's actual SAT scores. Although accuracy is sometimes intertwined with moral and legal issues, establishing accuracy is clearly a different endeavor than establishing morality or legality.

Sometimes, being accurate may be widely seen as offensive and immoral, especially if one is accurate when referring to others with intellectual, physical, or social limitations. This issue is particularly acute with respect to perceptions of groups. It is usually offensive to claim that "they," as a group, have less of some socially desirable attribute or engage less in some socially desirable behavior than do other groups. This chapter primarily addresses issues of accuracy, without regard to whether accuracy might be considered offensive or immoral by some people or by the legal system. Moral and legal issues are extremely important and will be touched on in a later section addressing political objections to accuracy research and the complex interrelationships of accuracy, bias, and discrimination. Legal and moral issues are, however, generally beyond the scope of this chapter (see, e.g., Dawes, 1988; and Gottfredson, 1994, for discussions of situations in which U.S. courts declared it illegal for employers to use the most valid information available for making decisions about hiring new employees).

V. Structure of This Review

The subsequent critical evaluation of objections to accuracy research is divided into three major sections: a review of the major conceptual objections to accuracy research, an evaluation of the methodological criticism

that it is inordinately difficult to identify criteria for assessing the accuracy of social judgment, and an evaluation of the claim that it is only possible to assess accuracy with statistically sophisticated and difficult componential analyses.

VI. What Are the Main Conceptual Objections to Accuracy Research? A Critical Evaluation of 50 Years of Criticisms

This section critically evaluates some of the most common substantive or theoretical objections to accuracy research (methodological and statistical objections, such as criteria and components, are considered in later sections). This includes claims that research on accuracy is less informative than research on error and bias because accuracy research provides little information about psychological processes; research on accuracy is unnecessary because research on social cognition already shows that social perception is generally inaccurate; research on accuracy is uninformative because it fails to address how perceivers explain others' behaviors and attributes; accuracy, itself, may not be very meaningful because that which is "accurately" perceived may merely reflect prior self-fulfilling prophecies; and stereotype accuracy research is unnecessary because even a stereotype that fits a group will inaccurately describe most members of that group. It also addresses a political objection to accuracy research: that accuracy research, and especially research on the accuracy of social stereotypes, is immoral because it does little more than justify or reify existing social inequalities.

A. "COGNITIVE PROCESSES ARE IMPORTANT, ERROR AND BIAS IS IMPORTANT, BUT ACCURACY IS NOT"

1. Process

Researchers studying cognitive processes have frequently ignored accuracy and have periodically explicitly repudiated it (see, e.g., Fiske & Taylor, 1991; Jones, 1985, 1986; Stangor, 1995). One reason may be that, in its most simplistic form, accuracy research seems to provide no information about process. Knowing only that a perceiver accurately judged some person or group does not tell us very much about that perceiver's thought processes.

This simplistic view of accuracy, however, does not characterize most accuracy research, nearly all of which does examine issues of social and cognitive processes. Brunswik's (1952) Lens Model, Cronbach's componential model,

Funder's (1995, 1999) Realistic Accuracy Model, Judd and Park's (1993) componential model for studying stereotype accuracy, Ickes (1997) work on empathic accuracy, Jussim's (1991) Reflection-Construction Model, and Kenny's (1994) Social Relations Model all address the processes by which people arrive at accurate or inaccurate judgments regarding others (several of these will be discussed in detail later in this chapter). Thus, the idea that research on accuracy is unimportant because it does not address issues of process is false because process issues are central to accuracy research.

2. Error and Bias

Error and bias may be considered more important than accuracy by some researchers for several reasons. One may be that errors and biases are widely viewed as revealing psychological processes. This is true (see, e.g., Funder, 1987; Kahneman & Tversky, 1973), but this then becomes the "process" reason for not studying accuracy, which, it has just been argued, is flawed because accuracy research also addresses process.

Another possibility is that bad and foolish behavior is inherently attention grabbing, and rational and reasonable behavior rarely is (Aronson, 1999; Krueger & Funder, 2004; Schneider, Hastorf, & Ellsworth, 1979). It is frequently surprising and dismaying to discover that we make systematic errors in certain types of situations. Identifying such errors is a potentially important first step toward correcting them and arriving at better judgments.

These reasons for emphasizing error and bias have validity. None, however, justify ignoring or dismissing accuracy research or the evidence of accuracy often available in studies of error and bias. Bad and foolish behavior may be attention grabbing, but presumably, scientists evaluate a phenomenon's importance by factors such as its frequency, power, and place within some wider theoretical context. For example, scientists typically consider car driving more dangerous than airplane flying, because far more people die per mile traveled in cars than in airplanes. That airplane crashes are far more attention grabbing is seen as an interesting social psychological phenomenon but not as a basis for emphasizing the dangers of flying. Similarly, the greater attention-grabbing power of error and bias does not constitute a scientific basis for considering them more important than accuracy.

Furthermore, it is only possible to correct errors if we know what the correct judgment should be and we have means of assessing whether future judgments are correct. Accuracy research is essential if one is interested in correcting systematic errors. Therefore, an interest in improving human judgment does not constitute a basis for considering error and bias research more valuable than accuracy research. Overall, therefore, although these

reasons for emphasizing error and bias have validity, they do not justify deemphasizing accuracy.

B. RESEARCH ON ACCURACY IS NOT NECESSARY BECAUSE THE SOCIAL COGNITION LITERATURE ALREADY SHOWS THAT SOCIAL PERCEPTION AND JUDGMENT ARE DOMINATED BY ERROR AND BIAS

Psychological research articles are filled with excellent experimental studies of cognitive processes that researchers interpret as indicating that bias, error, and self-fulfilling prophecy are likely to be common in daily life (e.g., Chen & Bargh, 1997; Fiske & Neuberg, 1990; Fiske & Taylor, 1991; Gilbert, 1995; Jost & Kruglanski, 2002; Kahneman et al., 1982; Nisbett & Ross, 1980; Stangor & McMillan, 1992). Therefore, another reason for dismissing or avoiding accuracy research may be that it is not seen as necessary. If errors and biases are identified, then, *ipso facto*, do we not know that people are inaccurate? In fact, however, we do not know this.

Conclusions regarding the power and prevalence of error and bias in daily life are only justifiable by research that examines the accuracy of people's judgments, preferably under realistic conditions. No matter how much researchers believe that the processes discovered in the lab should lead to bias and error in daily life, the only way to find out is by assessing the accuracy of those social perceptions. This occurs for at least two separable reasons: some biases can coexist with some degree of accuracy, and some biases enhance accuracy.

For example, consider the "false" consensus effect—the phenomenon whereby people overestimate the extent to which others agree with them. A simple example shows how this is not necessarily mutually exclusive with accuracy. Consider a simple dichotomous choice, say between candidates from two political parties. Assume that 75% of the people in this district belong to one party and 25% to the other party, that all people prefer the candidate from their own party and the existence of an extreme false consensus effect in this district—all people believe that their candidate receives the support of the majority. Of the people holding this belief, 75% will be correct (see also, Dawes, 1989).

Similar analyses showing that bias can either coexist with or enhance accuracy have been applied to expectancy effects, social stereotypes, and interpersonal perceptions (Brodt & Ross, 1998; Jussim, 1991; Kenny, 1994; Kenny & Acitelli, 2001; McCauley et al., 1980; Ryan, 1995, 2002). For example, stereotypes sometimes influences person perception judgments—a

pattern routinely interpreted as evidence of bias and source of discrimination (e.g., Aronson, 1999; Bodenhausen, 1988; Borgida, Rudman, & Manteufel, 1995; Darley & Gross, 1983; Fiske & Taylor, 1991; Fiske & Neuberg, 1990; Jones, 1986, 1990; Stangor, 1995). If the following two conditions are met, however, both Bayes's theorem and regression principles predict that such a bias will enhance the accuracy of those person perception judgments: individuating information is less than perfectly diagnostic, and the stereotype is itself an accurate characterization of the group (Jussim, 1991). Such predictions were later confirmed by research empirically demonstrating that the college students who relied most heavily on stereotypes regarding the types of students who choose to live in different residence halls (hippies, preppies, etc.) also reached the most accurate judgments regarding particular individual residents (Brodt & Ross, 1998).

As a consequence, even if experimental laboratory demonstrations of error and bias readily generalized to daily life judgments, perceptions, and decisions—which is not always clear—they do not inexorably lead to the conclusion that those judgments, perceptions, or decisions are inaccurate. If one wishes to make general claims about the extent to which social perception and judgment are dominated by inaccuracies, one cannot rely simply on error and bias research; one needs to assess the accuracy of those judgments.

C. ACCURACY OF EXPLANATIONS: "JUST BECAUSE SOME BELIEF ABOUT SOME PERSON OR GROUP IS CORRECT DOES NOT TELL US WHY OR HOW THE PERSON OR GROUP GOT THAT WAY"

One common criticism of accuracy research is that it fails to address explanations—either scientific ones or phenomenological ones—for how individuals or groups developed their characteristics. This criticism of accuracy research is implicit in perspectives indicating that accuracy cannot be studied or is meaningless because social processes and phenomena (e.g., discrimination, poverty) create the differences that are perceived (e.g., Claire & Fiske, 1998). It is similarly implicit in the common claim that stereotypes are flawed because they assume group differences are biological when they are not (see, e.g., critical reviews of the stereotype concept appearing in Ashmore & Del Boca, 1981; Brigham, 1971; Jussim, McCauley, & Lee, 1995; McCauley et al., 1980). Such perspectives indicate that demonstrating that people accurately perceive group differences is meaningless because it provides no information about either how people explain those differences or how those differences arose.

The claim that “demonstrating accuracy does not explain how or why those being accurately perceived got that way” is absolutely true; however, it also fails to threaten or undermine the viability, importance, or informativeness of accuracy research. This applies to most existing research on accuracy.

For example, friends have been found to be more empathically accurate than strangers (Stinson & Ickes, 1992). Empathic accuracy refers to inferring one’s partner’s conscious thoughts and feelings. Consider a Perceiver who believes a Target is feeling hostile. This “objection” focusing on the accuracy of explanations leads to at least four different questions: Is the Perceiver right? What is the Perceiver’s explanation for the Target’s hostility? If the Target is hostile, how did he or she get that way? And why does the Perceiver believe the Target is hostile?

Providing an answer to one question provides no information about the others. For example, establishing that the Perceiver is correct (the Target really is feeling hostile) tells us nothing about how the Perceiver explains the Target’s hostility. Maybe the Perceiver is a bigot who thinks that the Target’s ethnicity makes him or her prone to hostility. Maybe the Perceiver thinks the Target was mistreated as a child. Maybe the Perceiver thinks that the Target watches too many old Clint Eastwood movies.

Similarly, establishing that the Perceiver is correct tells us nothing about how the Target became hostile. Maybe there are genes for hostility, and the Target has them. Maybe he or she was abused as a child. Maybe the Target has been watching too many old Clint Eastwood movies.

One could attempt to establish the validity of the Perceiver’s explanation for the Target’s hostility by comparing it to the “true” reasons for the Target’s hostility, if they could be uncovered. Doing so would probably be a difficult task, but whole bodies of research have addressed sources of hostile and aggressive behavior (e.g., virtually every social psychological textbook has an entire chapter devoted to explaining aggression), so it would not be impossible. Assessing the validity of the Perceiver’s belief that the Target is hostile is simply a different endeavor than is assessing the validity of the Perceiver’s explanation for the Target’s hostility. That a particular study only focuses on assessing one type of accuracy does not fatally flaw such research—it only means that although considerable information may be provided regarding one type of accuracy (e.g., accuracy in perception of a feeling or trait), no information may be provided about another type of accuracy (e.g., accuracy in the explanation for that feeling or trait).

Furthermore, none of this necessarily explains how or why the Perceiver came to believe that the Target is hostile. This is a social and cognitive process question, and process is important, but it is not an accuracy question. The accuracy issue evaporates here, because we are no longer evaluating the

validity of the Perceiver’s judgments, expectations, or beliefs. With respect to evaluating the validity of the Perceiver’s belief that the Target was feeling hostile, it does not matter how the Perceiver explains the Target’s hostility, it does not matter how the Target came to feel hostile, and it does not matter how the Perceiver came to believe that the Target felt hostile.

This analysis is equally applicable to evaluating the accuracy of people’s beliefs about groups (stereotypes). No information about how people arrived at their beliefs about group differences, how they explain those differences, or how those differences actually emerged is provided by the now-abundant research showing that people’s beliefs in some differences between two groups are often partially or largely accurate (e.g., Brodt & Ross, 1998; Diekmann, Eagly, & Kulesa, 2002; Hall & Carter, 1999; Judd, Park, Ryan, Brauer, & Kraus, 1995; Jussim, Eccles, & Madon, 1996; Madon, Jussim, Keiper, Eccles, Smith, & Polumbo, 1998; McCauley & Stitt, 1978; McCauley & Thangavelu, 1991; Ryan, 1996, 2002; Ryan & Bogart, 2001; Swim, 1994). The lack of information about explanations constitutes a limitation to these studies. This limitation, however, does not threaten or undermine what they do show—considerable accuracy in people’s perceptions of differences between demographic groups or between individuals from different groups.

D. ACCURACY VS. SELF-FULFILLING PROPHECY: “IT IS NOT MEANINGFUL TO DISCUSS ‘ACCURACY’ IF WHAT IS BEING ‘ACCURATELY PERCEIVED’ DOES LITTLE MORE THAN REFLECT SELF-FULFILLING PROPHECIES”

This objection specifies a very particular process by which those being perceived accurately became that way—self-fulfilling prophecies. This review gives it separate consideration because numerous researchers have specifically stated or can be read as implying that accuracy is meaningless because that which is accurately perceived could result from self-fulfilling prophecies; many perspectives on expectancies enthusiastically embrace self-fulfilling prophecies but question accuracy, in part on the grounds that accuracy is confounded with self-fulfilling prophecy; the main type of “accuracy” some theoretical perspectives discuss is the specious form that results from self-fulfilling prophecies; and both accuracy and self-fulfilling prophecy involve a belief or expectation corresponding well with targets’ outcomes so that the potential confounding of the two is particularly salient or obvious (Claire & Fiske, 1998; Jones, 1986, 1990; Jost & Banaji, 1994; Olson, Roese, & Zanna, 1996; Snyder, 1984; Snyder & Stukas, 1998; Swann, 1984).

The logic underlying this objection seems to be the following:

1. Self-fulfilling prophecies occur.
2. Therefore, differences between targets may reflect effects of self-fulfilling prophecies.
3. If differences that are perceived reflect self-fulfilling prophecies to some unknown degree, attributing "accuracy" to those perceptions is, at best, unjustified, and at worst, reifies differences produced through social processes.

There is some truth to this argument. Differences that are accurately perceived at some point in time may reflect effects of prior self-fulfilling prophecies. Furthermore, the confounding of self-fulfilling prophecy and accuracy clearly would be a problem in any situation (e.g., daily life, lab research) in which it was not possible to distinguish these two very different reasons for why a perceiver's expectations might be confirmed. Simply showing that a perceiver's belief corresponds well with targets' actual attributes or behaviors, by itself, cannot distinguish accuracy from self-fulfilling prophecy. Researchers have, however, done far more than simply demonstrate correspondence between perceiver beliefs and targets' attributes and behaviors.

1. A Wide Array of Methodological and Statistical Techniques Exists for Distinguishing Accuracy from Self-Fulfilling Prophecy

One technique is to have people judge targets with whom they do not interact (e.g., by judging them from resumes, college records, photographs, etc.). People cannot create self-fulfilling prophecies among targets with whom they do not interact. Therefore, by ruling out self-fulfilling prophecy, such designs allow for an assessment of the accuracy of person perception judgments (for examples, see Archer & Akert, 1977; Brodt & Ross, 1998; Goldman & Lewis, 1977).

Other methods allow for the simultaneous assessment of accuracy and self-fulfilling prophecy. Although a detailed discussion of these methods is beyond the scope of this chapter, the core idea is simple: If perceivers' expectations are self-fulfilling, they should predict changes in targets' behavior or accomplishments over time. Theoretical models relying on structural equation techniques have been developed for distinguishing self-fulfilling prophecy from accuracy under naturalistic conditions (e.g., Jussim, 1991; Madon, Gyll, Spoth, Cross, & Hilbert, 2003; Trouilloud, Sarrazin, Martinek, & Guillet, 2002; West & Anderson, 1976; Williams, 1976).

Therefore, criticisms along the lines of "if accuracy is confounded with self-fulfilling prophecy, the meaning of accuracy is clouded" are true,

but only in the narrow literal sense of this conditional statement (i.e., if confounded, then clouded meaning). Numerous approaches have been developed that largely or completely unconfound the two.

2. "Prior Self-Fulfilling Prophecies May Influence That Which is 'Accurately' Perceived"

The self-fulfilling prophecy problem, however, does not go away so easily. Even if any particular study can rule out self-fulfilling prophecy in that particular study, the differences that are "accurately perceived" may still have resulted from prior self-fulfilling prophecies that occurred in interactions outside the study (either with other perceivers or with the same perceiver outside the study). Much like the prior criticisms of accuracy research, however, although this claim is valid (prior self-fulfilling prophecies may indeed influence that which is accurately perceived), it does not undermine the value, viability, or interpretation of accuracy. The next sections explain why.

3. The Unconfounding of Impressions and Predictions. I. The Perceiver Is Accurate Even if Self-Fulfilling Prophecies Resulting from Other Perceivers' Expectations Did Create Target Differences

Isn't it vacuous at best and misleading at worst to attribute "accuracy" to a belief that is true only because of prior self-fulfilling prophecies? When this rhetorical question is treated scientifically, rather than rhetorically, the answer is a clear "no." Understanding why requires understanding the difference between impressions and predictions.

If target behavior, accomplishments, and so forth predate perceiver beliefs about the target, causality can only flow in one direction: from target behavior to perceiver beliefs. Those perceiver beliefs may indeed become self-fulfilling, but only with respect to future target behaviors. In this chapter perceiver beliefs developed on the basis of prior target behaviors, accomplishments, and so on are referred to as "impressions"; and perceiver beliefs that might predict future target behaviors are referred to as "predictions." The importance of this distinction is illustrated with a composite, concrete example based on the type of data in numerous naturalistic studies of teacher expectations (e.g., Hoge & Butcher, 1984; Jussim et al., 1996; West & Anderson, 1976; Williams, 1976).

Consider Miss Smith, a sixth-grade teacher, beginning a new school year. In looking through the records of her new students, she discovers Donna, who received As last year and scored in the top 10% on a statewide standardized achievement test, and Mary, who received Cs and Ds and scored in

the bottom 30% of the same test. Miss Smith concludes that Donna has been a better student, a higher achiever, and knows the material better. Is there anything wrong, inappropriate, unjustified, confusing, confounding, uninterpretable, or vacuous about Miss Smith's evaluations? Is Miss Smith "blaming the victim?" Is she reifying differences between Donna's and Mary's ethnic groups?

These are rhetorical questions because it is obvious that Miss Smith's judgment is accurate, even if the difference between Donna and Mary resulted, in part, from prior self-fulfilling prophecies. Perhaps Donna's parents strongly encouraged her participation in intellectual activities, whereas Mary's parents did not. Perhaps Donna had a great teacher in fifth grade who inspired in her a commitment to educational achievement, and perhaps Mary had a teacher who was insulting and obnoxious and who discouraged her. These and other potential self-fulfilling prophecy explanations do not change the fact that here and now, Donna has a much better record than does Mary. Believing anything else would be starkly incorrect.

4. The Unconfounding of Impressions and Predictions. II. The Perceiver's Impressions Can Be Accurate Even if Self-Fulfilling Prophecies Resulting from the Same Perceiver's Expectations Did Create Target Differences

Again, the key issue here is time. If a swim instructor's expectations trigger a learning interaction sequence such that those expectations cause the student (Jacques) to become an unusually fast swimmer, those expectations are self-fulfilling (e.g., Trouilloud et al., 2002). But, once Jacques has completed the race in record time, how should the swim instructor perceive him? Would the swim instructor be most accurate if she perceived Jacques as a slow, awkward, incompetent swimmer? Or even as an average swimmer? Again, the answer is obvious. A "problem" arises only when we fail to distinguish between impressions and predictions (keeping in mind that today's impression can become tomorrow's prediction).

5. Self-Fulfilling Prophecies Occur but Do Not Invalidate Accuracy

Although self-fulfilling prophecies sometimes occur, one cannot assume that differences between any two people or groups result from self-fulfilling prophecies. Such a claim requires specific empirical justification (i.e., empirical evidence that self-fulfilling prophecies caused the differences among the particular targets being studied). Citation of a handful of dramatic self-fulfilling prophecy studies (e.g., Rosenthal & Jacobson, 1968; Snyder, Tanke, & Berscheid, 1977; Word, Zanna, & Cooper, 1974) does not

constitute adequate justification for a new claim that self-fulfilling prophecies caused differences among a new set of targets because so much other research shows that expectancy effects are far from inevitable, powerful, or permanent (see, e.g., meta-analyses and reviews by Brophy, 1983; Jussim, 1991; McNatt, 2000; Raudenbush, 1984; Rosenthal & Rubin, 1978). Thus, the claim that accuracy cannot be studied because prior self-fulfilling prophecies might influence that which is accurately perceived includes a core falsehood ("accuracy cannot be studied") enveloped in a good and valid point ("prior self-fulfilling prophecies might influence that which is accurately perceived"). Prior self-fulfilling prophecies might have influenced that which is perceived, but it does not mean accuracy cannot be studied.

E. STEREOTYPE ACCURACY AND LEVELS OF ANALYSIS

The following represents a distillation of a criticism of the notion of stereotype accuracy that has periodically appeared in the social psychological literature (e.g., Allport, 1955; American Psychological Association, 1991; Fiske, 1998; Hamilton, Sherman, & Ruvolo, 1990; Nelson, 2002; Stangor, 1995):

Even if it can be successfully shown that perceivers accurately judge two groups to differ on some attribute,

1. Perceivers cannot assume that their stereotypes of the group automatically fit all members of the group;
2. Perceivers cannot apply their belief about the group when judging individuals;
3. If perceivers do apply their belief about the group when judging individuals, they are likely to be wrong much of the time because few members perfectly fit the stereotype.

If all stereotypes are known to be largely inaccurate (as this logic suggests), the need to assess their accuracy would be rendered moot.

This criticism has some validity, but that validity depends, in part, on what this type of statement means. To the extent that the "perceivers cannot" statements represent moral injunctions, rather than statements about accuracy, they are beyond the scope of a consideration of the accuracy of social beliefs.

However, to the extent that "perceivers cannot" means "they would reach inaccurate judgments if they did," these arguments are a central focus of this chapter. This line of reasoning's suggestion, however, that all stereotypes are inaccurate because most members of a group fail to fit a stereotype is only partially justified. It is true that most members of a group will fail to

perfectly fit a stereotype. This, however, does not mean that the stereotype is inaccurate. To understand why requires understanding how this reasoning confounds at least three different levels of analysis and how considerably greater conceptual clarity can be brought to understanding stereotype accuracy by clearly distinguishing among these levels of analysis. Table I presents an analytic breakdown of different levels of analysis at which accuracy can be assessed.

1. Stereotypes as Perceptions of Populations

The first row in Table I refers to stereotypes—beliefs (or generalizations) about whole populations (typically, but not always, large demographic groups). The level at which one must measure the criterion for assessing the accuracy of beliefs about groups is the population that comprises that group. Claims about the characteristics of New Yorkers (or women or African Americans or librarians) should be compared with the characteristics of a representative sample or the whole population of New Yorkers (or women or African Americans or librarians). It is not possible to evaluate the accuracy of a belief about Asians in general by using as a criterion the characteristics of a particular Asian target. To do so would be equivalent to evaluating the claim that “Alaska is cold” by measuring the temperature at noon on July 4 in Anchorage.

Census figures, results from randomly selected samples, and meta-analyses of hundreds of studies have all been justifiably used as criteria against which to compare the accuracy of people’s stereotypes (e.g., Judd et al., 1995; McCauley & Stitt, 1978; Swim, 1994; although, how to measure criteria when assessing accuracy is discussed in detail later in this chapter). The evidence that has slowly accumulated regarding the validity of people’s beliefs about populations is that they are usually, though not always, at least moderately accurate, and that people are about as likely to underestimate as to overestimate real differences between population groups (see, e.g., Diekman et al., 2002; McCauley, 1995; Ryan, 2002). Such research, however, cannot and was never intended to evaluate the accuracy of people’s perceptions of individuals from different groups, which requires a level of analysis below that of whole populations.

2. Stereotypes and Person Perception: Accuracy and Bias in Judging Differences Between Individuals from Different Groups

The second row of Table I presents a second level of analysis for assessing accuracy—that of perceptions of differences between individuals belonging to different groups. Perceivers making claims about the differences between the Christians and atheists they know personally, or between their own sons

TABLE I
IDENTIFYING THE APPROPRIATE LEVEL OF ANALYSIS IN STUDIES OF SOCIAL PERCEPTUAL ACCURACY

Level of analysis	Social belief is a	Level of criteria for assessing accuracy of that social belief
<p>Population (This level assesses the accuracy of a stereotype about a group). <i>Research examples:</i> Judd et al., 1995; McCauley and Stitt, 1978; Swim, 1994</p>	<p>Stereotype Regarding an Entire Population <i>Examples:</i> 1. An introductory psychology student believes that White Americans are wealthier than African Americans. 2. A high school teacher believes that teenage boys are better at math than are teenage girls.</p>	<p>Population 1. Income of White Americans and African Americans in a nationally representative sample or in the U.S. Census. 2. Meta-analyses of hundreds of studies assessing sex differences in teenagers’ math performance.</p>
<p>Small Group (This level assesses the accuracy of beliefs about differences between specific individual targets belonging to different groups). <i>Research examples:</i> Brodz and Ross, 1998; Clarke and Campbell, 1955; Madon et al., 1998</p>	<p>Perception of Differences Between Specific Individual Members of Social Groups <i>Examples:</i> 1. An introductory psychology student sees little difference between the wealth of African American and White students in his class. 2. A high school teacher believes the girls in her class are doing better at math than are the boys in her class.</p>	<p>Small Groups 1. The wealth (net worth; yearly income) of the African American and White students in that student’s introductory psychology class. 2. Performance in class and on standardized tests of the boys and girls in this teacher’s class.</p>
<p>Individual (This level assesses accuracy in perceptions of individuals, not in perceptions of population or small group differences). <i>Research examples:</i> Funder, 1987; Jussim, 1989; Kenny, 1994</p>	<p>Person Perception <i>Examples:</i> 1. An introductory psychology student believes that Mary Anne is wealthier than Rashid who is wealthier than Lois. 2. A high school teacher believes that John is doing better at math than Bonita who is doing better than Lou.</p>	<p>Individuals 1. Mary Anne’s, Rashid’s, and Lois’s wealth. 2. John’s, Bonita’s, and Lou’s performance on math tests.</p>

and daughters, or between the police in two different towns, are all making claims about differences between small groups. They are not necessarily making claims about whole populations. A soccer coach might believe that the girls on her team play better than do the boys on her team without necessarily believing that girls are typically better soccer players than are boys. This level of analysis addresses the role of stereotypes in causing systematic inaccuracy in perceivers' judgments about individuals they know personally. Such claims occur at a different, smaller, level of analysis than do claims about differences between whole populations.

Assessing the accuracy of the perceived difference at this level of analysis must be accomplished by comparing the perceived mean difference between targets from differing groups to the actual mean difference. A handful of studies have addressed accuracy at this level of analysis. For example, Clarke and Campbell (1955) had students (African American and White) predict each other's performance on an upcoming test. White students' perceptions were a mix of accuracy and bias (slightly underestimating African American achievement), whereas the African-American students predicted the achievement of the African-American students quite accurately (they did not report results for each group's predictions of White students). Other research has assessed the accuracy of teachers' perceptions of differences between students from different demographic groups in their classes and the accuracy of college students' perceptions of individuals belonging to different college residence halls, and has found a similar pattern of high accuracy and small bias (Brodt & Ross, 1998; Jussim et al., 1996; Madon et al., 1998).

3. Person Perception

The third row of Table I presents a third level of analysis—the individual target.¹ At this level of analysis, most stereotype accuracy questions disappear. Accuracy in the perception of differences between individuals belonging to different groups can no longer be assessed. Without some comparison of differences in perceptions of groups (large or small), only accuracy in the judgment of individual targets can be assessed.

For example, many studies examine how well teachers' perceptions of their individual students' performance corresponds with those students' actual performance (Jussim & Harber, in press). How teachers arrives at

¹This corresponds to what is frequently called the dyadic level of analysis (e.g., Jussim, 1991; Kenny, 1994), because there is one perceiver and one target. However, the discussion of all three levels of analysis here has assumed a single perceiver, and that what is varying is the number and nature of the targets (population, small group, individual). Therefore, "person perception," is discussed here as occurring at the individual level of analysis.

their perceptions—whether through the use of stereotypes, shoe size, or acute insight—is not relevant at this level of analysis to evaluating the accuracy of their judgments of individual students. No matter how they arrived at their evaluation of a student, that evaluation either is or is not well-justified by that student's actual performance. Similarly, each of several perceivers might judge the personality attributes of each of several new acquaintances. Accuracy is determined by comparing the judgments to some criteria for measuring those acquaintances' personalities. Accuracy at this individual level of analysis is frequently at least moderately high (e.g., Funder, 1987; Jussim, 1991; Kenny, 1994).

Although stereotypes might influence those judgments, the role of such influence in accuracy can rarely be determined at this level of analysis. Instead, to determine whether stereotypes increase or reduce accuracy in judgments of individuals, researchers will usually need to group targets (by gender, ethnicity, or whatever stereotype one is studying) and then compare perceived group differences to actual group differences (i.e., at the intermediate level of analysis shown in Table I). For example, a business owner's evaluations of employees might correlate .6 with those employees' overall performance, indicating moderately high accuracy. Such accuracy tells us nothing, however, about whether the owner exaggerates differences between males' and females' job performance.

Few of the large number of studies comprising the stereotypes and person perception literature were designed to assess accuracy. Instead, most were intended to assess psychological processes involved in stereotyping. Thus, although they do not address accuracy, this limitation does not constitute a threat to the main purposes of such studies. It is nonetheless important to highlight their inability to address accuracy, because so much of the discussion of this literature states or implies that stereotypes typically lead to inaccurate judgments of individuals (American Psychological Association, 1991; Darley & Fazio, 1980; Fiske, 1998; Fiske & Neuberg, 1990; Jones, 1986, 1990; Jost & Kruglanski, 2002).

One reason for this common misinterpretation of this literature is that it often yields results showing that, even when targets have identical personal characteristics, accomplishments, or behaviors, people view targets differently, depending on whether they belong to one group or another. If two people (or people in two experimental conditions) hold different views of targets with identical personal characteristics, doesn't this imply that at least one of them must be wrong? Not necessarily.

First, few such studies included targets with real attributes against which the accuracy of perceiver judgments could be evaluated. In the absence of targets with real behaviors and attributes, it is often impossible to evaluate the accuracy of perceiver judgments. For example, there rarely are

standards against which to compare perceivers' behavioral predictions or trait inferences regarding fictitious targets invented by researchers.

Second, many researchers seem to assume that when targets' behavior is identical, perceivers reaching different conclusions regarding targets from different groups are inherently inaccurate. This assumption is true sometimes, but not always. For example, if two students get 8 of 10 math problems correct on a test, perceiving those students as correctly solving 8 of 10 problems is the only correct perception. In this situation, if two individuals or groups differ in their estimation of the number of problems correctly solved, one or both must be inaccurate.

However, if the judgment task involves inferring their general level of intelligence or predicting their future performance, it might not be accurate for the inferences and predictions to be identical. For example, if one student was in an enriched math class and the other in a remedial math class, the most accurate inferences and prediction might not be identical, despite the students' identical performances on this one test. Most likely, the students were placed in these different classes because they had dramatically different records of past math achievement. Even if the perceiver was not aware of those differences, they would typically be more accurate predicting that the student in the enriched class would perform more highly in the future than would the student in the remedial class (as compared to predicting their performances to be identical).

Last, in the rare cases in which experimental studies of the role of stereotypes in person perception include clear criteria for evaluating accuracy, almost none include no label or no stereotype control groups. It is possible that, although label A may increase accuracy relative to label B, any label increases accuracy compared to no label. To reach conclusions about whether relying on a stereotype or having a label to guide perception increases or decreases the accuracy of perceivers' judgments, of course, such control groups would be necessary. Although many perspectives suggest that reliance on stereotypes is both inaccurate and immoral (e.g., American Psychological Association, 1991; Devine, 1995; Fiske, 1998; Fiske & Neuberg, 1991; Jones, 1986; Stangor, 1995), the only study providing data capable of empirically comparing the accuracy of person perception judgments in the presence versus absence of a label (Cohen, 1981, experiment 2) found that having a stereotype label before exposure to a target increased accuracy.

4. *Stereotype Accuracy and Level of Analysis: Conclusion*

Claims suggesting that stereotypes are inaccurate because they do not apply to all individual members of a group (Allport, 1958; American Psychological Association, 1991; Fiske, 1998; Hamilton et al., 1990; Nelson,

2002; Stangor, 1995) are both true and false. The claim that stereotypes cannot possibly apply to all individual members of a group is completely true. The suggestion that this renders stereotypes inaccurate is, however, unjustified because it confounds two levels of analysis (population and individual). A claim about a population cannot be evaluated against the characteristics of an individual, or even against subsets of individuals. Consistency between the level of the perception and the level of the criterion must be maintained when assessing accuracy by comparing beliefs about populations (stereotypes) to characteristics of those population groups, beliefs about differences between small groups of individuals to the actual differences between those small groups of individuals, and beliefs about an individual to the characteristics of that individual.

5. *The One Exception: Absolutist Stereotypes*

Absolutist stereotypes—beliefs that all members of a group have some attribute—will indeed almost always be false, because there are almost always wide variations among individuals. A single exception invalidates an absolutist belief. Just as a belief that the temperature in all locations in Alaska is always below freezing will be disconfirmed by a single reading of 33 degrees Fahrenheit in Juneau on July 15 at 1 p.m., a belief that all Germans are efficient will be disconfirmed by discovery of a single inefficient German.

The vast accumulated empirical evidence on stereotypes, however, has yet to report a single person who holds absolutist stereotypes. Instead, the evidence indicates that most stereotypes are quantitative and probabilistic, not absolute (e.g., Judd et al., 1995; McCauley & Stitt, 1978; Swim, 1994). Probabilistic stereotypes, which permit many exceptions and wide variability, can only be evaluated by comparison to population-level criteria. People who hold absolutist stereotypes undoubtedly exist and probably comprise significant portions of extremist groups such as the Ku Klux Klan and neo-Nazis. Nonetheless, such people are atypical of the participants in most scientific research on stereotypes.

F. POLITICAL OBJECTIONS

Politics seems to lead to both explicit and implicit objections to accuracy research—especially research on the accuracy of stereotypes. First, people's political stances may lead them to be more likely to raise scientific concerns about research that they perceive as opposing their political positions than about research that supports their political stances (e.g., Lord, Ross, &

Lepper, 1979). This will almost never be stated quite so explicitly. No researcher will ever state "I am a liberal" (or "I am a conservative") or "Because I find this research politically offensive, I am going to work extra hard to come up with intellectual arguments against it." Furthermore, people may often not even be aware of how their politics influences their reactions to research.

Politics, however, seems to be a major basis for considering stereo type accuracy research to be a "problem" (e.g., Fiske, 1998; Stangor, 1995) while at the same time not considering research on, for example, automatic stereotyping, ingroup favoritism, ambivalent sexism, or subtle racism to be problems. These areas have their own substantial degree of theoretical and empirical controversy and complexity, so it would be difficult to make the case that stereotype accuracy research is more problematic than other areas on purely scientific grounds. Thus, political issues often remain implicit undercurrents underlying objections to accuracy research, rather than explicit statements. Occasionally, however, political objections to accuracy research have been explicitly articulated.

1. What Is the Political Criticism of Accuracy Research?

The main political objection is that accuracy research, at best, does not help anyone, and at worst, it helps exacerbate social inequalities. Specifically, accuracy research has been accused of doing little more than reifying or reflecting socially constructed injustices or assisting bigots and oppressors in their nasty goals (e.g., Claire & Fiske, 1998; Fiske, 1998; Jost & Banaji, 1994; Stangor, 1995). Before evaluating the validity of this criticism, why accuracy motivates this type of politically based criticism is briefly discussed.

2. Why Does Accuracy Arouse Political Motivations?

Accuracy runs against the grain of many social scientists' concern for helping alleviate social inequalities and injustices. Demonstrating that some aspects of many people's sex stereotypes are accurate (Diekmann et al., 2002; Hall & Carter, 1999; Swim, 1994) or that some aspects of many people's racial stereotypes are accurate (McCauley & Stitt, 1978; Ryan, 1996, 2002; Ryan & Bogart, 2001) does little to alleviate or explain injustices associated with sexism or racism. Worse yet, demonstrating social perceptual accuracy can be viewed not merely as documenting high acumen in perceiving individual and group differences but as reifying those differences.

To say that a belief that some child is unintelligent or is awkward on the basketball court or is socially inept is "accurate" has a feel of "blaming the victim." If the belief is accurate, then we cannot point to perceivers' errors, biases, misconceptions, egocentrism, or ethnocentrism as explanations for target difficulties. The unintelligent, unathletic, or socially awkward target, in these cases, really is flawed in some way. It is especially distasteful to suggest that some negative beliefs regarding large demographic groups (i.e., stereotypes) are valid. People who publicly declare that two groups differ in some societally valued attribute (intelligence, motivation, propensity for alcoholism or crime, morality, etc.) run the risk of being accused of being an "-ist" (racist, sexist, classist, etc.) or, at minimum, of holding beliefs that do little more than justify existing status and hierarchy arrangements (e.g., Jost & Banaji, 1994; Sidanius & Pratto, 1999).

In contrast, an emphasis on bias and error implies a benevolent and egalitarian concern with injustice. Such an emphasis implies that the emphazier believes that so-called "real" differences between groups do not result from any actual attributes of members of those groups (their cultures, their religions, their histories, their economic conditions, their geography, their practices, their politics, their genetic predispositions)—they result solely or primarily from the oppressive effects of others' misbegotten beliefs.

In addition, an emphasis on cognitive errors and biases provides a clear villain—the person committing the error or bias. It also points to a relatively straightforward way to ameliorate some social inequities—change the source of the bias (people's beliefs, expectations, processing, motivations, etc.). Whereas belief change is not necessarily an easy thing to accomplish, it is much easier than changing the mean level of competence among a large portion of the population.

In contrast, if the belief about group differences is accurate, then changing the situation may require work on a grand scale. To make groups more equal, we must upgrade one group's competence (we could also downgrade the higher-status-groups' competence, which may appeal to some people with deeply felt resentments, but which most scientists would agree is not the optimal solution). Doing so may require years of education, training programs, mentoring, and the like. This is both much more labor intensive and typically not under the purview of most social psychological activities.

Furthermore, a series of studies on interracial stereotypes in the United States have converged on the conclusion that many Whites have adopted an "egalitarian" ideology that, as part and parcel of its emphasis on intergroup equality, actively denies differences between groups (Judd et al., 1995; Ryan, 1996, 2002; Wolsko, Park, Judd, & Wittenbrink, 2000). This denial-of-difference ideology is not shared by many African Americans, who have adopted a "multicultural" ideology that involves acknowledging both the

positive and negative differences between groups (see also Lee et al., 1995; Ryan, 2002 for discussions of the similarities between multiculturalism and stereotype accuracy). Nor is this denial-of-difference ideology shared by many researchers outside the United States—despite active programs of research on stereotypes and prejudice in Australia, Canada, and Europe, the most vigorous critiques of the concept of accuracy, especially when applied to stereotypes, have been produced by American researchers (e.g., Claire & Fiske, 1998; Fiske, 1998; Jones, 1985, 1986; Stangor, 1995). In contrast, for example, an Australian team, in a broad and sweeping review of research on stereotypes, characterized the social cognition emphasis on stereotype inaccuracy as “mythic” (Oakes, Haslam, & Turner, 1994).

Regardless, researchers (whether White or not, and whether in the United States or not) who share the type of denial-of-difference egalitarian ideology uncovered in the research by Ryan and Judd and their colleagues may be offended by research on stereotype accuracy because it implies the potential existence of bona fide differences between groups. This may also help explain the relatively greater political appeal to many social psychologists and other social scientists of error, bias, and process compared to accuracy, especially with respect to stereotypes.

3. Are the Political Criticisms Justified?

Does accuracy research blame the victim and justify and exacerbate inequalities, as suggested by the politically motivated criticisms? Social and political judgments have the potential to influence all sorts of research, especially in the social sciences (Herrnstein & Murray, 1994; Jacoby & Glauberman, 1995; Redding, 2001). In this sense, political analyses of accuracy research are justified, but no more than they are in other areas of research. Accuracy research is neither more nor less subject to such influences than other social science research.

The idea that accuracy research somehow reifies, creates, or justifies social inequalities and injustices is simply false—at least if the criterion is what researchers studying accuracy have actually written. I know of no articles that provide or review evidence of accuracy that state, imply, or even hint that accurately perceived differences between individuals or groups reflect immutable, permanent, or essential features of those people (see, e.g., Brophy, 1983; Eagly & Diekmann, 1997; Funder, 1987, 1995; Ickes, 1993, 1997; Judd & Park, 1993; Jussim, 1991; Kenny, 1994; Lee et al., 1995; McCauley et al., 1980; Ryan, 1995, 2002). Regardless, as in most social science fields, strong steps can be taken to considerably reduce political bias from entering into consideration of issues involving accuracy (discussed later in

this chapter; see also Funder, 1987, 1995; Judd & Park, 1993; Kenny, 1994; Ryan, 2002).

Whether it is scientifically and politically more appropriate to deny and ignore group differences or to understand and acknowledge such differences is clearly a contested and controversial issue. For both scientific reasons (discussed throughout this chapter) and political reasons (discussed next), accuracy research involves understanding and acknowledging such differences.

4. Accuracy Research As A Tool in the Alleviation of Social Problems

Even if the sole purpose of research were to alleviate social problems, wouldn't we want to know whether people's beliefs about groups and their individual members (i.e., stereotypes) are accurate? We should want to know for several reasons. First, if we think we are curing a social disease by changing people's inaccurate, biasing, or misbegotten beliefs about groups, our efforts will be sorely misplaced if their perceptions of groups and individuals are already accurate. Second, if some beliefs are widely inaccurate and some are reasonably accurate (as is likely the case), only research directly and empirically assessing the accuracy of stereotypes could possibly tell us which beliefs and whose beliefs need to be changed through social interventions.

Furthermore, we need to be able to assess and understand accuracy to improve the quality of people's judgments and evaluations. Only after determining that some people hold highly inaccurate beliefs would it be reasonable to begin work on changing those beliefs.

Work on changing inaccurate beliefs itself would only be useful if it were conducted after we knew how to lead people to arrive at more accurate judgments. Of course, there will be no way to assess our success at leading people to adopt more accurate beliefs unless we have techniques to assess accuracy. By understanding what leads people astray and what leads them to accurate judgments, we will be much more capable of harnessing those factors that lead to accurate judgments and, therefore, reduce social problems resulting from inaccurate beliefs.

5. Accuracy Can Lead To Discrimination

Although most of the political criticisms of accuracy and accuracy research are not justified, social perceptual accuracy can indeed lead to unequal distribution of valued resources to different groups and to discrimination. The perspective taken in this chapter is that it is more valuable to understand the social and psychological processes that might lead accuracy to cause discrimination than to simply deny the existence of

accuracy or to suppress accuracy research. At least two classes of such situations are distinguishable: situations in which accurately perceived group differences lead to justifiable differences in group outcomes, and situations in which accurately perceived group differences lead to unjustified outcomes for at least some individuals. Each class is discussed next.

a. Accurately Perceived Bona Fide Differences Between Individuals from Different Groups Can Lead to Group Differences in Outcomes. People with low incomes, for example, will typically not be eligible to receive as large a mortgage from a bank as are people with higher incomes. If bank loan officers accurately judge others' wealth and income, they will offer larger mortgages to wealthier people. Regardless of whether one considers this evidence of discrimination against poor people, demographic groups often differ in their wealth. On average, in the United States, Asians, Whites, Latinos, Native Americans, and African Americans rank from highest to lowest, respectively, in yearly household income (U.S. Census, 2002). If bank loan officers are good at judging wealth, and even if they are completely unprejudiced, Asians will be more likely to be approved for larger loans than will be African Americans.

Similarly, many middle-class jobs require a college education. In the United States, a higher proportion of Whites than Latinos complete college. Therefore, even if race and ethnicity-based discrimination evaporated, employers who accurately evaluate applicants' level of education would provide proportionately more of those jobs to Whites. Accurately perceived differences in skills, competencies, accomplishments, and so forth between individuals belonging to different groups can sometimes produce mean group differences in valued societal outcomes (jobs, college admissions, etc.).

b. Accuracy Leads to Bona Fide Discrimination. It is also possible, however, that accurately perceived group differences can simultaneously lead perceivers to be as accurate as possible in their judgment of individuals and to unfairly discriminate against many members of certain groups. Because this issue is politically charged, a nonsocial example is used first to illustrate the basic principles.

A U.S. penny is almost, but not quite, perfectly balanced—when flipped in the air, it will come down as heads slightly more frequently than tails. If one wished to successfully predict the outcome of a coin toss, one would maximize one's accuracy by predicting heads 100% of the time. Such a person could be justifiably characterized as having an "extreme heads bias," but the predictions would also be as correct as possible.

Such a situation characterizes U.S. car insurance policy rates. Young people, and especially young males, are legally charged far more for car insurance than are middle-aged adults. This is because young people,

and especially young males, are actuarially more likely to have accidents. Although undoubtedly true at the group level, this policy clearly discriminates against individual young adults who drive safely.

A structurally similar situation, but one in which the discrimination is illegal in the United States, is one in which a bank uses race/ethnicity as a basis for making loans. Because credit history and past income are not 100% diagnostic of future ability to pay, however, in an actuarial (or Bayes's theorem) sense, it might be more accurate to predict that an Asian with a mixed credit history will be more likely to pay off a loan in the future than will a Native American with an equally mixed credit history. Nonetheless, instituting a policy that denied loans to Native Americans with credit histories comparable to those of Asian Americans who received loans would clearly constitute illegal discrimination.

The potential for bias to enhance accuracy and then to produce discrimination is probably one of the least well understood (or perhaps merely least discussed) processes and set of interlocking phenomena in the social sciences. Bias is frequently seen as implying inaccuracy; accuracy is frequently assumed to be antithetical to bias and discrimination. Such a view is, however, not only seriously flawed, but may account for some of the resistance to and criticisms of accuracy, and especially stereotype accuracy, within social psychology. To the extent that researchers view accuracy as conflicting with their egalitarian goals of alleviating social problems, it is understandable that they may be tempted to denigrate accuracy research. But understanding that bias sometimes enhances accuracy, that accurately perceived group differences in inputs may lead to group differences in outcomes, or that biased and discriminatory predictions may be as accurate as possible under some circumstances seems crucial with respect to the scientific analysis of intergroup relations and social policy.

The goals of accuracy and social justice sometimes conflict. It seems unlikely that deciding how to reconcile those goals will be best accomplished by merely denying accuracy. It seems more likely that a thoughtful understanding of the ways in which accuracy, bias, and discrimination are intertwined, and not necessarily mutually exclusive, could contribute to informed and durable policy decisions regarding issues such as civil rights legislation, antidiscrimination laws, and affirmative action. Social scientists and lawmakers will be in a much stronger intellectual position to weigh the relative costs and benefits of accuracy and egalitarianism when making policy if they have a more complete understanding of those costs and benefits than if they merely deny lay accuracy or attempt to suppress research on accuracy.

6. Conceptual Objections to Accuracy Research: Conclusions

Within social psychology, accuracy research has had a turbulent and controversial history, which probably helps explain why such little research on accuracy was performed from about 1955 to 1985. This, in turn, may help explain the minimal discussion of accuracy in most major reviews of social cognition and social perception phenomena, such as stereotypes, expectancies, schemas, categorization, attribution, and person perception (see, e.g., the 1985, 1996, and 1998 editions of the *Handbook of Social Psychology*, and the last 30 years of *Annual Review* chapters on these topics).

This section has reviewed many common conceptual objections to accuracy research. It has suggested that many of these objections have considerable validity, in the sense that the criticisms may be true and in that they often raise interesting and important questions. Accuracy in perceptions of an attribute really is different than accuracy in explanations for that attribute. Prior self-fulfilling prophecies might explain some differences between targets that are accurately perceived. Despite whatever validity they might have, however, this critical evaluation of the common criticisms leads to the conclusion that none constituted serious threats to the value or interpretation of existing research on social perceptual accuracy.

One important and common objection to accuracy research has not yet been discussed—the so-called criterion “problem.” The selection and identification of appropriate criteria for evaluating accuracy has been frequently characterized as a problem sufficiently severe to call the viability of accuracy research into question (e.g., Fiske, 1998; Jones, 1985, 1990; Snyder & Stukas, 1998; Stangor, 1995). Next, therefore, this review will critically evaluate whether such issues are more complex or problematic than those facing most areas of psychological research.

VII. How Can One Establish Criteria for Evaluating the Accuracy of Social Perception?

Only gods know Absolute Truth. The goal of accuracy research is not to establish Absolute Truths. Instead, the criteria against which social beliefs should be evaluated are similar to the criteria against which psychological hypotheses are evaluated—truths with a small “t.”

A. TRUTH WITH A SMALL “T”

Truth with a small “t” is evidence, preferably from a variety of sources, indicating that some belief is valid (true, warranted, justified, etc.). In short, the solution to the problem of partially fuzzy, intangible constructs is to establish their validity through multiple methods and approaches (e.g., Campbell & Stanley, 1963; Cook & Campbell, 1979; Cronbach & Meehl, 1955). Truth with a small “t” is typically established via rigorous methodological variations of the “duck” test (Block, 1993; Funder, 1995). If it looks, walks, acts, and sounds like a duck, although the possibility remains open that it really is an antelope, sport utility vehicle, alien visitation, or bacterial growth, it is most likely a duck. A person is, of course, more likely to confuse a duck with a goose or a cormorant than with the Statue of Liberty or a praying mantis. Even if that goose is believed to be a duck, however, *probabilistic realism* further assumes that it would eventually be possible to receive information that corrects this faulty view.

The same analysis applies to all the fuzzy attributes studied by psychologists. Does Voter A have a positive attitude toward the current president of the United States? We might ask A whether she supports the current president, voted for him in the last election, likes him, would vote for him in an election held today, and belongs to the same political party as the president. If Voter A says she does support and like the president, voted for him last time, would vote for him again, and belongs to his political party, classifying her as having a positive attitude toward the president does not seem unreasonable. A similar analysis could be applied to mental and emotional states, self-esteem, personality, expectations, schemas, intelligence, stereotypes, prejudice, motivation, and indeed, the whole pantheon of constructs studied by psychologists. The myriad of papers on these and similar hypothetical, intangible constructs are written as if the authors believe these constructs are real. If the constructs filling the journals and texts are “real enough” for psychologists to study in their scientific research, why are they not “real enough” to constitute criteria for evaluating the accuracy of social perception, social beliefs, and social judgments?

B. CONSTRUCT VALIDITY AND SOCIAL PERCEPTUAL ACCURACY

There are only two logically coherent possibilities with respect to the issues of construct validity and criteria for social perceptual accuracy. One possibility is that the constructs routinely studied by psychologists have no

reality, so that the articles relying on and reporting them are meaningless and it makes no sense to use meaningless criteria in research on accuracy. This perspective, though logically coherent, is far too extreme and has few proponents [except perhaps some radical social constructivist work that characterizes mainstream social science research as, in essence, story-telling (e.g., Gergen, 1985)].

A second possibility is that the constructs routinely studied by psychologists are real enough, so that the articles relying on and reporting them are meaningful. It would, therefore, make sense to use those same criteria when studying accuracy. This is the perspective advocated here. Of course, even this perspective permits the possibility of there actually being no good criteria for certain types of social perceptions, judgments, or beliefs. If scientific research has yet to develop a valid method for measuring a construct that constitutes the content of some perception, judgment or belief, then it would actually and indeed literally be the case that there is no valid measure against which to assess the accuracy of that perception, judgment, or belief. Given the wide variety of measures available for assessing all sorts of psychological constructs, however, such situations are likely to be the exception, rather than the rule.

1. A Logically Incoherent Possibility: The Constructs Psychologists Study Are Real, But There Are No Good Criteria for Assessing Accuracy

Articles that take for granted the validity of numerous social psychological constructs, such as attributions, perceptions, stereotypes, identity, attitudes, and so on, sometimes claim that establishing criteria for assessing accuracy is problematic (e.g., Fiske, 1998; Jones, 1985; Stangor, 1995). In general, it is logically incoherent to assume that the constructs studied by psychologists are real but to characterize accuracy criteria as problematic. One cannot claim that attitudes, personality, behavior, emotions, and so forth can be studied and then claim that we have no criteria against which to evaluate people's beliefs about others' attitudes, personalities, behaviors, or emotions.

2. Self-Fulfilling Prophecies, Accuracy, and the Criterion "Problem" Double Standard

An extraordinary double standard regarding criteria issues has emerged within the social cognition tradition. Many researchers have embraced and emphasized the power of self-fulfilling prophecies as a mechanism by which social beliefs, and especially stereotypes, create objective social realities (e.g., Claire & Fiske, 1998; Fiske & Taylor, 1991; Jones, 1986, 1990; Jost &

Kruglanski, 2002; Miller & Turnbull, 1986; Snyder, 1984, 1992). Many of these same researchers, however, have expressed doubts or concerns about the viability of identifying objective scientific criteria for establishing accuracy (see also Fiske, 1998; Kruglanski, 1989; Jones, 1985; Snyder & Stukas, 1998).

This is peculiarly ironic because, although the processes by which a perceiver's beliefs become true are different for accuracy and self-fulfilling prophecy, the criteria for establishing their trueness must be identical. To establish either accuracy or self-fulfilling prophecy, one must show that a belief corresponds to some criterion. There is a difference only in the process—in how the correspondence comes about—but the criteria for establishing whether any particular belief is true must be just as good (or bad) as the same criteria, regardless of whether they are used in self-fulfilling prophecy research or in accuracy research.

Researchers are logically coherent if they accept the criteria used to establish correspondence between perception and reality in both self-fulfilling prophecy and accuracy research, a position that is well-justified. A position suggesting that "both accuracy research and self-fulfilling prophecy research should be dismissed" because of criteria problems is internally consistent and logically coherent. This position is so extreme, however, that no articles advocate it. It would be, however, logically incoherent to accept the criteria used to establish correspondence between perception and reality in the service of self-fulfilling prophecy research while rejecting the same criteria when used to establish correspondence between perception and reality in the service of accuracy research.

C. WHAT IS AND IS NOT INCLUDED IN THIS SECTION ON CRITERIA

This next section identifies and describes four broad classes of criteria: objective criteria, behavior, agreement with others, and agreement with targets. Each type of criterion is defined and distinguished from other types of criteria. When a certain type of criterion has been controversial, those controversies are discussed and critically evaluated. Finally, the viability of each type of criterion is also critically evaluated. This next section, however, only addresses how to identify plausible criteria. It does not address how best to use those criteria for estimating accuracy (this is addressed in a subsequent section on componential and other approaches for assessing accuracy).

The term "criteria" here is not used in the general sense of "how to assess accuracy." Even the limited aspects of "how to assess accuracy" that are addressed in this review involve the issues raised throughout this entire

chapter (not any one section) and cannot possibly reflect all the issues and details involved in assessing accuracy in all the ways and contexts in which it has been assessed (see e.g., Ambady & Rosenthal, 1992; Brunswik, 1952; Cronbach, 1955; Funder, 1995; Ickes, 1993, 1997; Judd & Park, 1993; Jussim, 1991; Kenny, 1994; McCauley et al., 1980 for such details). Thus, the term "criteria" is used quite narrowly herein to refer to the measures, variables, or constructs that might be used as standards against which to evaluate the validity of social perceptions, judgments, and beliefs. As shall be seen, however, even the issues surrounding criteria defined narrowly are both complex and controversial.

D. OBJECTIVE CRITERIA

Criteria are objective when that which is being judged is assessed in a standardized manner that is independent of the perceiver's judgment. Many interesting and important social phenomena involve simple, clear, objective criteria. A tennis player wins or loses the point. A salesman makes or loses the sale. An applicant is either accepted into or rejected from a college.

There is nothing the least bit difficult or "problematic" about the criteria for predictions regarding these events. Ninety-nine times out of 100, or more, there will be no controversy on whether the ball was in or out. In the event that one player hits a clear and obvious winner, we all agree that it was a winner, but the criterion for evaluating whether or not it was a winner is not our agreement. We agree because we all saw the ball clearly land in and because the second player did not even hit the ball. Agreement is a result of accuracy, not a criterion for establishing accuracy. Even John McEnroe, perhaps the most argumentative professional tennis player of all time, at his most argumentative only argued a few calls per match—and a tennis match includes hundreds of shots.

All sorts of outcomes occur independent of individual perceivers' beliefs, predictions, or expectations (e.g., election outcomes, returns on investments, the achievements and accomplishments of our friends and acquaintances, etc.). These come up all the time in daily life, and the criterion issue is rarely a "problem."

1. One-Shot Hit or Miss: The Prototypical Case

It is perhaps easiest to understand the elegant simplicity in identifying and using objective criteria by starting with a class of cases, referred to here as "one-shot hit or miss." This means there is a single perceiver making a single judgment about an objective single attribute or behavior of a single target.

Although finding out whether a single person makes an accurate or inaccurate single prediction is not usually particularly psychologically interesting, by virtue of eliminating a slew of methodological, conceptual, and statistical issues irrelevant to the criterion issue, the one-shot hit or miss situation with objective criteria, though oversimplified, is particularly informative.

Despite many social psychologists' emphasis on pervasive ambiguity in social reality (e.g., Gilbert, 1995; Jones, 1990; Ross & Nisbett, 1991), many aspects of human social behavior, including but not restricted to sports, relationships, investments, academics, and politics, are objective. If I predict that my friend's marriage won't last 2 years, the criterion is similarly clear and objective—they either do or do not get divorced within 2 years. If Margaret predicted that the Republicans would win a majority of Senate seats in 2002, she was right. If, in early 2000, Bob predicted that the stock market was in the midst of a speculative bubble that was destined to crash and put all his money into money markets and bonds, he clearly made the right decision.

The processes by which Bob, Margaret, and I arrived at our beliefs—our propensities to be self-serving, expectancy-confirming, reliant on faulty heuristics, ideologically motivated, or dispositionist in our attributions—are all irrelevant to evaluating whether we are right or wrong. The one-shot hit or miss situation makes it clear that objective criteria exist for many personally and socially significant beliefs, expectations, and predictions, and accuracy can be assessed independent of process.

2. Overall Levels of Accuracy

One-shot hit or miss situations, however, reveal few, if any, important psychological principles and rarely have much bearing on any psychological theory or hypothesis. Usually, therefore, research on accuracy requires investigating many people making one or more judgments or predictions about one—and sometimes many—other person(s).

Although statistical procedures for assessing accuracy sometimes become more difficult in such situations, the criterion issue is no more difficult than in the one-shot hit or miss situation. If a criterion is appropriate for assessing the accuracy of some prediction or judgment in the one-shot hit or miss situation, it is equally and identically appropriate for assessing accuracy when there are multiple judges, judgments, or targets. Dollars in real estate sales, numbers of college acceptances and rejections, and length of a romantic relationship are all exactly identical in their appropriateness as criteria (for judgments of real estate sales success, predictions of college admissions success, and length of relationship predictions, respectively), regardless of

whether one person judges one other person or 1000 people judge 1000 other people.

For example, Archer and Akert (1977) developed a test of nonverbal sensitivity that involved assessing people's ability to accurately identify objective aspects of people, their experiences, or their relationships. In two of their test questions, people were exposed to two women playing with a baby and talking and two men who had just finished a one-on-one basketball game.

Participants' task was to correctly identify the mother of the baby and the winner of the game. In the actual test, there is a series of such vignettes and the more questions answered correctly, the more accurately people are at judging others on this test. This is simple and objective; there is no criterion problem here.

The criterion issue is also a nonproblem in much of the deception-detection literature. That is, one question psychologists have asked is "How good are people at detecting when others lie?" For example, Ekman and Friesen (1969, 1974) had targets view either a pleasant film or a horrible gory one. All had to inform perceivers that they had just seen a lovely, pleasant film. The research question was how often people can tell when others are liars. Again, there is no criterion problem.

In studies of some aspects of stereotypes, one can also compare people's perceptions to objective data. In one classic study, McCauley and Stitt (1978) compared people's beliefs about differences between African Americans and other Americans to U.S. Census data (e.g., in size of family and likelihood of completing high school). In others, people's beliefs about sex differences (e.g., in aggression) were compared to the results from meta-analyses assessing sex differences (Hall & Carter, 1999; Swim, 1994). Although even the U.S. Census and meta-analyses do not necessarily reveal Absolute Truth, they do provide strong sources of evidence.

With this type of data, one can establish overall levels of accuracy, error, and bias. Do people err on the side of believing what others say? Then Ekman and Friesen's studies should have yielded results showing that people underestimate deception (it did, but their research also typically showed that people do better than chance at accurately detecting deception). Do people exaggerate differences between demographic groups? Then McCauley and Stitt (1978) and Swim (1994) should have found people consistently overestimating the differences between different groups (they did not—people tended to be accurate and, when wrong, to underestimate real differences). These are important and interesting accuracy questions, and the availability of clear, objective criteria in these cases is not problematic.

3. Independent and Standardized but Not Universally Persuasive Objective Criteria

Not all people may agree that certain objective criteria are good ones. Such agreement might be irrelevant regarding a target's height or marital status, but it becomes much more relevant when estimating a target's extraversion or intelligence via a personality questionnaire or standardized IQ test. Is the personality questionnaire a good one? Is it reliable? Valid? IQ tests, in particular, have a long and controversial history (e.g., Gould, 1978; Herrnstein & Murray, 1994; Neisser, Boodoo, Bouchard, Boykin, Brody, Ceci, Halpern, Loehlin, Perloff, Sternberg, & Urbina, 1996), as has the measurement of enduring personality traits (see Funder, 2001, for a review).

To the extent that some people do not find such tests credible, they are likely to discredit or dismiss research on accuracy using such criteria. Thus, use of objective but controversial criteria can be viewed as boiling down to agreement (if one agrees with the criteria, the study assesses accuracy; if one does not agree with the criteria, it does not—see Kruglanski, 1989), and socially and politically, this is probably how things work. People who do not accept a study's criteria most likely will not accept its conclusions (whether on accuracy or any other social science topic).

Often, however, the reverse may happen: People who do not like a study's conclusions will come up with arguments against the appropriateness of its criteria (Lord, Ross, & Lepper, 1979). Indeed, people are often extremely good at engaging in the intellectual contortions necessary to maintain a cherished viewpoint even in the face of clear evidence that they are wrong (Aronson, 1999; Festinger, Riecken, & Schachter, 1956). The fact that some people do not like or respect some criterion, of course, does not mean that it cannot be used. It only means that one needs to make a credible case for the appropriateness of that criteria in one's study.

For example, standardized test scores have often been used as criteria in studies of the extent to which teacher expectations for students were self-fulfilling prophecies or accurate (see edited volumes by Brophy, 1998; Dusek, 1985; see reviews by Brophy, 1983; Brophy & Good, 1974; Jussim, 1986; Spitz, 1999). If teachers are good at identifying which students are smarter, then their expectations should predict but not influence students' standardized test scores. If teacher expectations do predict standardized test scores without causing them, any real or imagined weaknesses in standardized tests would seem largely irrelevant to interpreting this result.

Use of imperfect criteria will generally lead researchers to underestimate people's accuracy. If so, in sharp contrast to attempts to dismiss the viability of accuracy research because of imperfect criteria, this means that people are

NOTICE

This material may be protected by
Copyright Law (Title 17 U.S. code)

probably even more accurate than indicated by the evidence emerging from any particular study using imperfect criteria. This is obviously the case when using a criterion measure with less than perfect reliability. Unreliability in measurement artificially lowers correlations between judgments and criteria, so that the greater the unreliability, the higher the actual accuracy.

Consider a case in which people's perceptions of others' extraversion correlate .3 with a self-report extraversion scale. If there is less-than-perfect reliability in measurement of both perception and criterion, then the true correlation will be higher than .3. If, for example, the reliability of both scales was .9, the true correlation between perception and criterion would be .33; if the reliability of both scales was .75, the true correlation between perception and criterion would be .4 (Carmines & Zeller, 1979).

This would often be equally true with a scale of imperfect or partial validity. Perhaps a particular psychological measure only captures a partial aspect of some attribute. People's judgments, perceptions, and so forth, if they reflect more aspects of the attribute than the measure being used, may be more accurate than indicated by the correlation between judgment and criterion. Consider an assessment of people's accuracy in judging how big a target is. The criterion, however, is height. Height, of course, is only one aspect of size (the others being width and depth). If people use height, width, and depth to judge size, the correlation of their judgments with the (imperfect, partial) criterion will be too low. That means their accuracy could be even higher than indicated by that criterion.

Similarly, consider the use of an IQ test as criterion for evaluating the accuracy of judgments of intelligence. If a particular IQ test primarily taps verbal intelligence and people's judgments include, in addition to estimations of verbal intelligence, creativity, social skill, political savvy, wit, and common sense, if people's judgments closely correspond to actual intelligence defined in this multifaceted way, the correlation between judgments and criterion will be artificially low. People might be more accurate than indicated by the correlation (correspondence) between judgment and criterion. An identical analysis applies to using any measure, such as a standardized personality inventory, as a criterion for assessing accuracy in judgments of personality traits.

Of course, just because such mismatches may lower the observed correspondence between perception and criterion, one cannot infer high accuracy from low correlations, but the argument that any particular criterion in any particular study is "bad" (unreliable, low validity) means that whatever evidence on accuracy is obtained probably often represents a lower bound on accuracy. This, in turn, means that people may actually be more accurate than indicated by the empirical evidence of accuracy obtained in that study. Just as imperfect criteria do not preclude the possibility of

objective, social scientific research on all sorts of topics (aggression, identity, achievement, social class, etc.), nor do they preclude the possibility of accuracy research.

Some criteria are subject to debate and controversy. The extent to which nearly all measures of social science constructs have been validated or well established has limitations, thereby inflicting imperfections on any research using any measure of any construct. Researchers should almost always carefully evaluate the flaws and limitations of their research, including their use of measures. However, this issue is no more a problem for accuracy research than for any other type of research.

E. BEHAVIOR

1. *What Is Behavior?*

In the broadest sense, anything a person (or any organism) does is behavior. In this sense, answers on a standardized test constitute behavior. Similarly, agreement among judges usually constitutes agreement concerning either targets' behavior or inferences about their attributes based on their behavior (Kruglanski, 1989). Even completely unobservable thinking, feeling, or daydreaming can be considered behavior. Indeed, transmission or inhibition of neurotransmitters can be considered behavior at the level of the individual cell.

None of this is what is meant by behavior here. For the social perception contexts addressed in this chapter, behavior is considered to be observable action, not unobservable thought processes, cell electrochemical processes, or underlying attributes. Counting how frequently Natasha smiled is related to but different from determining whether she is thinking happy thoughts or how sociable she is. Smiles are visible. Similarly, measuring how much time Jorge spends on some task is related to but different from evaluating Jorge's motivation. Again, time on task is an observable behavior, not an underlying attribute.

Similarly, "answers on a standardized test" is not what most social scientists mean when they use the term "behavior." Such tests are generally designed to measure some attribute that has relevance beyond the testing situation. Intelligence test scores should predict school outcomes, work outcomes, and many life outcomes. Scores on a personality scale measuring hostility should predict levels of angry and aggressive behavior in all sorts of interpersonal contexts. Although standardized test scores could be considered one class of behavior, it is both possible and useful to distinguish between standardized tests and overt actions.

2. Does Behavior Boil Down to Agreement?

Some have argued that it does (e.g., Kruglanski, 1989), and for good reasons. At least in most research contexts, a common way to measure human behavior is by having independent judges observe and code the behavior. However, there is a difference between the need for people to observe and record behavior and the need for people to interpret behavior. Behavior itself is observable; personality, beliefs, attitudes, motivation, competence, intelligence, and so forth may be inferred from behavior, but they are not directly observable. Observers recording nonverbal behaviors (e.g., smiling, speech dysfluencies, etc.), time spent on some task (e.g., proportion of time spent talking in a dyadic conversation; time spent trying to solve an impossible anagram before giving up; time it takes to walk down a hallway, etc.), task choice (how many advanced math courses does a person take? Do they apply to college? Number of parties attended, etc.) are all recording objective aspects of behavior.

Around the fringes, there may be room for interpretation—Was Fred smiling, or was the corner of his mouth just twitching a bit? And it is very useful to have multiple observers record behavior in order to resolve these fringe or ambiguous situations. When that happens, judgments of behavior do boil down to agreement. Nonetheless, when behavior rather than underlying attribute is being recorded, such fringe or ambiguous situations are generally likely to be the exception rather than the rule. Thus, there is an important distinction between using observers to record what a person does and using observers to interpret and evaluate the meaning of what a person does. Behavior is what a person does.

Many of the classic programs of research in social psychology have studied phenomena with behavioral criteria, such as research on the attitude-behavior relationship (Eagly & Chaiken, 1993), self-efficacy (Bandura, 1977), behavioral confirmation (Snyder, 1984), obedience and conformity (Asch, 1955; Milgram, 1974), and helping in an emergency (e.g., Latane & Darley, 1970). I know of no critical work that takes these areas to task for their use of behavioral criteria as somehow “problematic” or even for “merely studying agreement.” The rationale for believing that behavior can be identified when one studies such classic social psychology topics as these, but not when one studies accuracy, has never been explicitly stated in any prior article criticizing or rejecting accuracy research or arguing that accuracy boils down to nothing more than agreement (e.g., Claire & Fiske, 1998; Fiske, 1998; Jones, 1986, 1990; Kruglanski, 1989; Stangor, 1995). It remains a challenge, therefore, for scientists arguing that it is somehow problematic to use behavior as a criterion for studying accuracy to explain why the use of behavioral criteria is not equally problematic in other areas of psychology.

F. AGREEMENT WITH OTHER PERCEIVERS

Agreement, at first glance, appears to be a poor criterion for accuracy. Agreeing that the world is flat does not make it flat. Thus, agreement, like other criteria, is almost always imperfect.

But agreement cannot be so readily dismissed. If both of us are accurate, we must also agree. If a baseball player hits a home run, and we both saw it, we will both agree that he hit a home run. If Hanna is a brilliant student, and we both recognize her brilliance, we will both agree that she is brilliant. Accuracy causes agreement, although agreement does not necessarily imply accuracy. Furthermore, if we disagree, at least one of us must be wrong (at least if we are talking about the same thing).²

This analysis indicates that accuracy typically increases agreement. In probabilistic or correlational terms, as accuracy among multiple perceivers increases, so should their agreement most of the time. In fact, this is one of the keys to understanding why agreement among multiple perceivers is often a good, if imperfect, index of accuracy. Correlations work two ways: If A is positively correlated with B, then it is equally appropriate to claim that as A goes up so does B as it is to claim that as B goes up so does A. If accuracy and agreement among multiple perceivers will typically be positively correlated, then, as their agreement goes up, so will accuracy most of the time. Thus, agreement is an imperfect but probabilistic indicator of accuracy.

In fact, however, even “agreement” itself is a multifaceted construct. Agreement with whom? Experts? Other people? The targets themselves? In fact, all of these types of agreement may, under different circumstances, constitute better or worse criteria for accuracy. The next section, therefore, describes and reviews the types of agreement criteria most commonly used in accuracy research and identifies the major potential limitations of each.

1. Agreement with Experts

Perceivers' judgments or expectations can be compared to those of experts. For example, personality or psychopathology judgments could be compared with those of professional experts, such as clinical psychologists or psychiatrists; ratings of students' aggressiveness or academic motivation

²We do need to be careful in considering what constitutes “the same thing.” You and I could disagree about how pleasant, extroverted, or intelligent Fred is. If Fred acts differently with you from with me, then we both could be right (Swann, 1984). So if we are talking about Fred in general, both of us may be partially right and partially wrong, because perhaps there is no “Fred in general”—there is only Fred interacting differently with different people.

might be compared with those of the students' teachers. For some types of research, nonprofessionals could also be considered "experts" in the sense that they have some sort of unique access to knowledge about the target. Thus, spouses, romantically involved couples, close friends, and coworkers could be used as experts with whom to compare perceivers' judgments of targets.

2. Agreement with Experts' Models

Sometimes, experts, such as statisticians, decision-making theorists, or psychologists, have developed formal models for what constitutes the most appropriate way to arrive at a judgment. Accounting for regression to the mean, appropriate use of base rates, and use of consensus, consistency, and distinctiveness in arriving at attributions all constitute formal models of the appropriateness of social judgment (see, e.g., Dawes, 1979; Kahneman et al., 1982; Kelley, 1967; Nisbett & Ross, 1980). Thus, judgments or decisions based on such models may be used as criteria against which to evaluate the accuracy of laypeople's judgments or decisions.

3. Agreement with Independent Judges

If a social interaction is recorded, for example, by videotape, transcription, and so on, or if it can be monitored (e.g., through intercoms, computers, or one-way mirrors), people other than the perceiver can be asked to evaluate the target. These people are referred to here as "independent" judges because they were not involved in the interaction between perceiver and target; thus, they are independent of the perceiver. Often in such situations, these independent judges may only be exposed to the targets' responses to further minimize any effects of the perceiver's verbal or nonverbal behavior on the independent judges (e.g., Goldman & Lewis, 1977; Word et al., 1974). Thus, the perceptions of independent judges may be used as criteria against which to evaluate the accuracy of the perceiver's person perception judgments.

4. Agreement with Nonindependent Judges

Sometimes the judgments of people who do interact with the target, either simultaneously while interacting with the perceiver or in other contexts, may be used as criteria for evaluating the perceiver's accuracy. In such cases, all perceivers' judgments of targets essentially constitute the criteria for one another. Indeed, Kenny's (1994) social relations model, which addresses numerous aspects of agreement, accuracy, and social perception, requires

multiple people to interact with and rate one another. Nonindependent judges may be strangers, acquaintances, or experts.

5. Limitations to Agreement with Other Perceivers

The perspective of probabilistic realism indicates that nearly all criteria will be imperfect to some degree. Agreement is no exception. The fundamental limitation to use of agreement is, of course, that everyone can be wrong.

a. Even Experts May Be Wrong. Indeed, sometimes, expert judgments or predictions may be little better than those of lay people (Cronbach, 1955). Psychiatrists, for example, predicted that hardly any teachers would give 450 volts worth of shock to learners providing wrong answers on a test in Milgram's (1974) studies of obedience (even though half or more often did).

In addition, experts may be subject to their own biases. Happy romantic couples, for example, may see each other in an overly positive, almost idealized way (Murray, Holmes, & Griffin, 1996). The doctors and staff at psychiatric hospitals at least sometimes misinterpret patient boredom or hostility as psychopathology (Rosenhan, 1973). Teachers' expectations sometimes color their interpretations of students' performance (Williams, 1976). Thus, expert judgments are susceptible to a variety of potential imperfections. They are probably best used as a criterion for accuracy when prior research has provided strong evidence for their validity.

b. Experts' Models May Be Contradictory. Expert models often have a formal logical rigor or statistical/mathematical basis that gives them immediate credibility. Nonetheless, they may not always be correct. This will be demonstrated below using two examples in which different expert models make opposing predictions (meaning that both cannot possibly be right): one involving the stock market and the other involving stereotypes.

i. Efficient markets theory versus regression to the mean. Efficient markets theory (the dominant view of the stock market among economists) states, in essence, that because whatever is known is already factored into stock prices, it is impossible to consistently obtain returns that beat the overall market (Malkiel, 1999). The stock market, according to this view, is essentially a random walk (the future is inherently unpredictable) with an upward overall trajectory. The practical implication is that most people who do not plan to spend their investment dollars any time soon will receive the greatest return by buying a stock index fund that reflects the entire stock market and holding through (unpredictable) ups and downs. One will most likely get oneself into trouble (reduce one's overall return) if one tries to "time" the market (buy low, sell high) or select individual stocks to beat the market averages.

Efficient markets theory, however, can be viewed as clashing with one of the main claims of statistical experts. The statistical idea of regression to the average means that extreme values are likely to return over time to the overall mean. This could be interpreted to mean that if stock valuations are unusually high, one is more likely to lose money and less likely to make money (going forward) than usual; and if they are unusually low, one is more likely to make money and less likely to lose money (going forward) than usual. If one enters the market when valuations are unusually low, therefore, and exits when they are unusually high, one should be able to beat the overall market (by receiving the unusually large gains that follow low valuations and avoiding the severe losses following high valuations).

ii. *Bayesian inference versus stereotypes "biasing" person perception.* A similar contradiction occurs in expert models regarding use of stereotypes. Many social psychological theories and perspectives on stereotypes, much current cultural discourse, and many laws and judicial decisions state or implicitly assume that people are acting rationally and appropriately only when they judge others solely and entirely on the basis of those others' personal characteristics, rather than their group memberships. This view argues or implies that people act in a biased, prejudiced, or irrational manner when they allow their stereotypes to influence their judgments of individuals (e.g., Borgida et al., 1995; Brewer, 1988a; Darley & Fazio, 1980; Dunning & Sherman, 1997; Fiske & Neuberg, 1990; Jones, 1986, 1990; Myers, 1999; Nelson, 2002).

According to widely accepted principles of probability (e.g., Bayes's theorem, see e.g., Kahneman & Tversky, 1973), regression (e.g., Jussim, 1991), and philosophy of science (e.g., Krueger & Funder, 2004; Meehl, 1990), however, base rates, prior beliefs, and expectations should often influence people's interpretations of new evidence (see McCauley et al., 1980 for an analysis focusing specifically on stereotypes). These principles suggest that if perceivers are not sure how tall a person is, they will, on average, be right more often if they estimate any particular male to be a few inches taller than any particular female than if they estimate them to be exactly equal (i.e., use, rather than discard, their sex stereotype regarding height). Even if perceivers find out that both a male and female target are avid basketball players (identical individuating information), the best guess, according to these principles, is still that the male is taller than the female. Only if the perceivers find out that both are 6 feet tall, which is perfectly diagnostic of height, should they not apply their belief that men are usually taller than women.

This set of alternative expert models has created the following amusing and ironic state of affairs within the stereotypes literature. Specifically, people are accused of irrationality if they ignore their stereotypes when judging individual targets because they are supposed to use base rates

when judging others under uncertainty (e.g., Kahneman & Tversky, 1973; Locksley, Borgida, Brekke, & Hepburn, 1980); and more typically, they are accused of error and bias if they use their stereotypes when judging individual targets (e.g., American Psychological Association, 1991; Brewer, 1988a; Darley & Gross, 1983; Dunning & Sherman, 1997; Fiske, 1998; Fiske & Neuberg, 1990; Jones, 1986, 1990; Nelson, 2002). These and other discussions of the role of stereotypes in person perception assume that people are acting in an irrational, biased, or prejudicial manner if they perceive mean differences (on whatever attribute or behavior is being judged) between individuals belonging to different social groups, holding individuating information constant (Bodenhausen, 1988; Borgida et al., 1995; Darley & Fazio, 1980; Hamilton et al., 1990; Miller & Turnbull, 1986). Relying exclusively on individuating information and ignoring stereotypes/base rates is often held up almost as a moral ideal to which people should strive.

It cannot be normatively appropriate to both use and ignore prior expectations. Thus, the expert assumption that people should judge others solely on the basis of individuating information, which has emerged from much of the stereotype literature, conflicts with the claim that prior expectations and beliefs should influence judgments (in the absence of perfectly diagnostic individuating information) that have emerged from the cognitive judgment and decision-making literature.

In general, explicitly stating one's model of rationality is typical of the judgment and decision-making literature (e.g., Kahneman et al., 1982) but atypical of the stereotyping literature. The few exceptions in the stereotype literature (e.g., Krueger, Hasman, & Acevedo, 2003; Locksley et al., 1980; McCauley et al., 1980) use Bayes's theorem as their model of rationality. Obviously, by adopting the same model that underlies much of the judgment and decision-making literature, these perspectives avoid the conflict between expert models concerning stereotypes.

Apparently, however, there are social and political risks in scientists relying on Bayes's theorem when attempting to understand accuracy, error, and bias in judgments of individuals. When McCauley et al. (1995) argued that a Bayesian analysis indicates that reliance on accurate beliefs about groups should increase, not decrease, the accuracy of person perception judgments in the absence of perfectly diagnostic individuating information, they were accused of "disagreeing with civil rights law" (Fiske, 1998, p. 381). Such an accusation confounds moral and legal issues with accuracy; if valid, represents another situation in which it may be illegal to be accurate; and highlights the role of politics in some criticisms of stereotype accuracy research.

Regardless, discussions emphasizing the power of stereotypes to bias judgments (e.g., Aronson, 1999; Borgida et al., 1995; Darley & Fazio, 1980;

Devine, 1995; Dunning & Sherman, 1997; Fiske, 1998; Fiske & Neuberg, 1990; Fiske & Taylor, 1991; Jones, 1986, 1990; Jost & Banaji, 1994; Stangor, 1995; Stangor & McMillan, 1992) have rarely grappled with the implications of Bayes's theorem and regression principles for understanding the role of stereotypes in enhancing accuracy in social perception. Considerably more conceptual clarity could be brought to research on stereotypes and person perception if researchers wishing to reach conclusions about error and bias explicitly stated and justified how they define rationality, both conceptually and operationally within their empirical research, and what type of pattern of person perception judgments would constitute reasonable, justified, appropriate, or accurate judgments under the circumstances being studied. In the absence of such explicit justifications, it will remain too easy for researchers to continue to interpret diametrically opposed judgments—those relying on stereotypes and those ignoring stereotypes—as both reflecting faulty and erroneous judgment processes.

G. AGREEMENT WITH THE TARGET

Agreement with targets' self-descriptions can and have been used as a criterion for assessing accuracy. Two broad types of targets' self-descriptions are distinguished here. The term "self-reports of behavior" is used to refer to the actions targets claim they have engaged in. Examples might be how many glasses of alcohol they consumed yesterday, how much time they spend exercising each week, how often they argue with their spouse, and how much sleep they get each night. In contrast, the term "self-perceptions" is used to include targets' attitudes, beliefs, feelings, and evaluations of themselves, their characteristics, and their accomplishments. Self-perceptions might include self-esteem, self-perceptions regarding personality traits (independent, assertive, extraverted, etc.), feelings, political positions, self-evaluations of academic or athletic performance, and so on. Self-perceptions, therefore, generally involve unobservable, underlying attributes of some type, whereas self-reports of behavior involve overt, observable actions.

1. Agreement with Targets' Self-Reports of Behavior

Targets' self-reports of behavior can be used as a criterion against which to evaluate perceivers' judgments. Social reality typically constrains bias (e.g., Jussim, 1991; Kunda, 1990), so that self-reports regarding specific and objective behaviors may be less likely to be tainted by self-serving,

defensive, or impression management biases than are self-reports regarding vague or ambiguous attributes.

To get concrete, a fraternity member's response to "How many alcoholic drinks have you had in the last week?" may be less likely to be biased than is his response to questions such as "How often did you get drunk last week?" Although bias may emerge regarding even the most objective of behaviors, there is a lot more room for interpretation in the "drunk" question ("well, I did drink two sixes of beer, but I never really got drunk") than in the "drinks" question. Similarly, responses to "How much money did you donate to charity last month?" are more likely to be constrained by reality than are responses to "How generous are you?" Again, bias is always possible, but even a highly biased person is not likely to interpret buying a pizza as donating to charity, although people may vary a great deal on whether they consider sharing the pizza with a friend as a hallmark of great generosity.

2. Agreement with Targets' Self-Perceptions

Despite their imperfections, targets' own self-perceptions regarding underlying or ambiguous attributes, such as personality characteristics or dispositions, can often be used as a criterion (see, e.g., Judd & Park, 1993; Judd et al., 1995; Ryan, 1995). As usual with probabilistic realism, the issue is not whether such a criterion is perfect, because no criteria are perfect. The question is whether the specific self-perceptions being used as a criterion are likely to reflect what the target is like.

Indeed, there are theoretical reasons for expecting self-perceptions to be good criteria, at least sometimes. People have access to much more information about some of their experiences, inner states, relationships, and so on than do outsiders. Indeed, at least some research has shown that people's memory for both randomly sampled and particularly striking thoughts and actions, up to 5 months after they occurred, is excellent (e.g., Barclay & DeCooke, 1988; Brewer, 1988b).

Because bias can exist side by side with accuracy (e.g., Jussim, 1991), even when biases taint self-perceptions, they do not necessarily eliminate their validity (see, e.g., Block, 1965; Funder, 1995). If so, they can be used in the same manner as any imperfect criteria. Self-perceptions of academic ability, for example, usually correspond highly with indicators of academic achievement, such as grades and standardized test scores (e.g., Eccles & Wigfield, 1985; Felson, 1984). Attributes that are more readily observable, such as sociability or extraversion, tend to generate high agreement between self-reports and observers' ratings (e.g., Funder, 1995; Kenny, 1994).

Furthermore, considerable validity evidence has been obtained regarding numerous self-report scales of personality characteristics and political attitudes (Robinson, Shaver, & Wrightsman, 1991, 1999).

More highly validated self-perceptions are, obviously, better criteria than less well validated self-perceptions. If well-validated by prior research, however, self-report scales cannot (or at least should not) be summarily dismissed as "merely" self-report, at least not without refuting the preexisting validity evidence for the scale. In the absence of clear validity evidence, however, self-reports must be interpreted with caution and due regard to the topic. Self-reports of height, broccoli-eating, and interest in pornography fall on a continuum of least to most likely to be contaminated by social desirability biases.

3. Limitations to Agreement with Targets' Self-Reports and Self-Perceptions

Targets, of course, are imperfect themselves. Although they may have unique access to certain types of information (personal experiences, feelings, etc.), many people are subject to both motivated and unmotivated errors and biases (Kunda, 1990; Nisbett & Ross, 1980). Sometimes, people do not have access to their own cognitive processes or to some of their implicit beliefs and attitudes (Nisbett & Wilson, 1977; Wilson, Lindsey, & Schooler, 2000). Memory is imperfect and potentially subject to all sorts of biases (e.g., Koriat et al., 2000; Stangor & McMillan, 1992), so that self-reports may often be imperfect records of behavior.

Researchers, however, have also developed a slew of methods for improving the accuracy of self-reports (e.g., event sampling, daily diary methods, etc.). Although even these types of methods may not completely eliminate error and bias, they have become widely used because of their demonstrated validity for many types of self-reports (e.g., Hedges, Jandorf, & Stone, 1985; Rääkkönen, Matthews, Flory, Owens, & Gump, 1999).

The problem of motivated biases may often (though not always) be greater when using self-perceptions (rather than self-reports of behavior) as a criterion, in part because of the common tendency for most people to view themselves in self-serving ways (Myers, 1999; Taylor & Brown, 1988). In addition, many people may either lie outright or slant their responses in such a manner as to present themselves in as socially desirable, moral, and competent a manner as possible (Paulhus, 1991, 2002). Although this means that the overall average level for some self-perception may often be too favorable, this may not prevent self-perceptions from being a good correlational criterion for assessing the accuracy of perceivers' beliefs.

For example, Bertha may think she is a great athlete and Nyesha may think she is a good athlete. Both may be overestimates (Bertha may only be pretty good and Nyesha may be pretty average), but if their degree of self-inflation is similar, it may be true that Bertha is more athletic than Nyesha. So a coach who views Bertha as more athletic than Nyesha would be correct (and the coach's views would correlate well with Bertha and Nyesha's self-perceptions). Thus, even biased target self-perceptions may, under many circumstances, constitute good, if imperfect, criteria for assessing the accuracy of perceivers. In practice, although a default rejection of self-perceptions as criteria is not justified, researchers considering the use of particular self-perceptions need to thoughtfully consider their limitations and develop a convincing case that, in the context under study, they are likely to be good criteria.

H. CRITERIA: CONCLUSIONS

Probabilistic realism occasionally provides a gold standard for establishing accuracy. A real estate salesman either did or did not sell the house. Except in rare cases, that is not a matter of opinion. The absence of such a gold standard, however, does not justify the conclusion that the criteria available to assess accuracy are so fundamentally flawed or their interpretation so ultimately ambiguous as to cloud the meaning of accuracy research. Establishing that a social belief or perception is accurate is much like establishing validity in social science research.

The strongest and clearest evidence regarding accuracy comes from research that typically uses multiple measures and methods to establish the accuracy of social perception (e.g., Cronbach, 1955; Funder, 1987, 1995; Kenny, 1994). This does not mean that research examining accuracy using a single method or criterion is uninformative. Such research, however, may be less informative than research using multiple methods or criteria (except when that single criterion itself has previously been validated using multiple criteria!).

Despite scientists' inability to achieve absolute 100% certainty, probabilistic realism, truth with a small "t," and the principles of construct validity provide clear ways of identifying criteria against which researchers can assess the accuracy of social perception. Once one has a good set of criteria, assessing accuracy would appear to be simple—just evaluate how well the perceptions correspond with the criteria. In fact, however, this turns out to be considerably more complex than it seems at first glance, and the next section explains why.

VIII. Must All Research on Accuracy Assess Components?

A stopped clock is right twice each day. That does not make it a good clock. What does this have to do with social perception? More than it seems. Consider the following example: A wife successfully predicts that her husband will come home at least half an hour later than he said he would come home. It looks like this wife knows her husband pretty well, doesn't it? Not necessarily. Maybe she always predicts that her husband will be late. Maybe she always predicts everyone will be late. It could even be worse than that: Maybe she always predicts all people will always do bad things and have bad attributes.

Even though she might have happened to have been right that one time, she could not necessarily be considered a particularly astute judge of her husband. One could think of her prediction as stemming from several sources or components: her overall tendency to think well (or poorly) of people, her overall tendency to predict that people will be late (over and above her general tendency to think well/poorly of people), her overall tendency to think that her husband will be late (over and above her general tendency to think well of people and that people will be late), and her specific tendency to predict that her husband will come home later than he claimed this particular time (over and above her tendency to think well of people; to predict that they, in general, will be late; and to think her husband will run late). Each component of her prediction can be accurate to some degree, and each contributes to her overall likelihood of being accurate (across lots of judgments or predictions).

A. COMPONENTIAL APPROACHES TO THE STUDY OF ACCURACY

This type of thinking inspired Cronbach's (1955) classic review and at least two other more recent perspectives (Judd & Park, 1993; Kenny, 1994), all of which identified several processes contributing to social perception and that argued that accuracy needs to be separated into different components reflecting these different processes. Next, therefore, a brief overview of each of these three componential approaches to the study of accuracy is presented.

1. Cronbach's Components

Cronbach's (1955) analysis involves computing an accuracy score for each perceiver judging a set of targets on a set of traits. Thus, there is a separate score for each perceiver. In this situation, both social judgment and social

reality can be divided according to effects for targets and traits (Kenny, 1994):

$$\text{Social Judgment} = \text{Constant(elevation)} + \text{Target main effect (differential elevation)} + \text{Trait main effect (stereotype accuracy)} + \text{Target} \times \text{Trait interaction (differential accuracy)}. \quad (1)$$

The constant (which Cronbach called "elevation") is simply the grand mean of all judgments, the target main effect (which Cronbach called "differential elevation") refers to mean differences in how each target is judged (over all traits), the trait main effect (which Cronbach called "stereotype accuracy") refers to mean differences in how each trait is judged (over all targets),³ and the target \times trait interaction (which Cronbach called "differential accuracy") refers to how each target is rated on each trait after subtracting out the other three effects.

Social judgments are then compared to criteria, which are broken down into the same sets of components. Accuracy, then, refers to how well each component of the judgment corresponds to the identical component on the criterion. If, for example, a student tends to see other people through rose-colored glasses, her elevation score will be higher than the elevation score on the criterion. If she sees Professor Smith as a better teacher than Professor Jones and the criteria also indicate that Professor Smith is a better teacher than Professor Jones, her differential elevation accuracy will be high (even if seeing the world through her rose-colored glasses leads her to overestimate both of them).

If, when rating a group of professors, she sees absent-mindedness as more common than is organization, when in fact few professors are absent-minded but many are well-organized, her stereotype accuracy will be low. And if she rates Jones as more organized than Smith (after subtracting

³Cronbach (1955) probably referred to the trait main effect as "stereotype accuracy" because it refers to how perceivers seem to view "people in general" on this particular trait, which could be considered a type of stereotype. This was unfortunate, however, because Cronbach's term is potentially confusing. It means something quite different from what most people usually think of as social stereotypes regarding, for example, race, class, sex, and so on. "Trait accuracy" would, in my opinion, have been a clearer label. In this chapter therefore, I use the term "stereotype accuracy" to refer to two very different phenomena, depending on the context. With one exception, I always use the term "stereotype accuracy" to refer to the validity of people's beliefs about the attributes and behaviors of groups and their individual members. The one exception occurs in sections (such as this one) discussing Cronbach's components. In the context of discussions of Cronbach's components, the term "stereotype accuracy" refers to the validity of people's perceptions of the level of a trait among the particular targets being rated. This use of the term is explained in more detail in the text discussing the meaning of each of Cronbach's components.

out elevation, differential elevation, and stereotype accuracy), and Jones really is more organized than Smith, her differential accuracy will be high [see Kenny (1994, pp. 117–121) for a fully explicated mathematical example of Cronbach's components involving three perceivers, three targets, and three traits].

2. Kenny's Social Relations Model

The Social Relations Model (SRM) is a componential model that is related to, but differs from, Cronbach's (1955) in several important ways (see, e.g., Kenny, 1994; Kenny & Albright, 1987, for detailed discussions of similarities and differences). First, it is intended to be a broad and general model for assessing many different aspects of social perception, of which accuracy is only one. Second, research using SRM typically focuses on perceptions regarding one trait at a time, rather than the multiplicity of traits addressed by Cronbach's components. Third, however, it also typically focuses on several perceivers, rather than the one perceiver (at a time) that was the focus of Cronbach's analysis. Thus, SRM research might perform one analysis to find out how accurately Dave, Charles, and Bella perceive each other's intelligence and another to find out how accurately they perceive each other's friendliness.

Whereas Cronbach partitioned the judgment into target, trait, and target \times trait components, SRM partitions judgment into target, perceiver, and target \times perceiver components. Thus, the SRM equation is:

$$\begin{aligned} \text{Each Social Judgment} = & \text{Constant (elevation) +} \\ & \text{Perceiver main effect (perceiver effect) +} \\ & \text{Target main effect (target effect) + Perceiver} \times \\ & \text{Target interaction (relationship effect).} \quad (2) \end{aligned}$$

The equation can be computed separately for each judgment provided by each perceiver, as shown by this example (from Kenny, 1994, p. 130):

$$\begin{aligned} \text{AI's judgment of Bob} = & \text{Constant +} \\ & \text{AI's perceiver effect +} \\ & \text{Bob's target effect +} \\ & \text{AI's relationship with Bob.} \quad (2a) \end{aligned}$$

The constant is the grand mean of all judgments, the perceiver main effect is the mean difference between perceivers (over all targets), the target main effect is the mean difference between targets (over all perceivers), and the perceiver \times target interaction constitutes the unique manner in which a

particular perceiver judges a particular target after subtracting out the constant, the perceiver main effect, and the target main effect.

In contrast to Cronbach's (1955) components, however, SRM has a substantially different equation for the criterion:

$$\begin{aligned} \text{Criterion/Social Reality/Behavior} = & \text{Constant +} \\ & \text{Partner effect + Actor effect +} \\ & \text{Partner} \times \text{Actor Interaction} \quad (3) \end{aligned}$$

This equation can be computed separately for each actor (also from Kenny, 1994, p. 130):

$$\begin{aligned} \text{Bob's behavior with AI} = & \text{Constant +} \\ & \text{Bob's Actor Effect + AI's Partner effect +} \\ & \text{Bob's relationship effect with AI} \quad (3a) \end{aligned}$$

The constant is the mean level of the criterion; the actor effect refers to mean differences between actors (averaging over all partners); the partner effect refers to mean differences in how all actors, on average, behave with each partner; and the partner \times actor interaction refers to the unique manner in which a particular actor behaves with a particular partner after subtracting out the constant, partner effect, and actor effect.

To assess accuracy, each component of the judgment is compared to the corresponding component on the criterion. Overall, perceivers may see a high level of intelligence in their partners. If, on average, the actors really are pretty smart, elevation accuracy will be high (the constants will be similar). If, on average, AI sees his partners as much smarter than they really are, AI's perceiver accuracy will not be high (comparison of AI's perceiver effect [Eq. 2a] to AI's partner effect [Eq. 3a]). If, on average, people see Bob as smarter than the others, and he really is, Bob's target accuracy (called "generalized accuracy" in SRM) will be high (comparison of Bob's target effect, i.e., how he is usually perceived [Eq. 2a], with Bob's actor effect, i.e., what he is usually like [Eq. 3a]). If, on average, AI sees Bob as particularly smart (after subtracting out the constant, AI's perceiver effect, and Bob's target effect from the judgment) and Bob really does act particularly smart when interacting with AI (e.g., as might occur if AI's perception of Bob is self-fulfilling), then AI's dyadic accuracy will be high⁴ (comparison of the relationship effects in Eqs. 2a and 3a).

⁴Although I do not consider self-fulfilling beliefs to be accurate, they could be considered a form of accuracy within SRM.

3. Judd and Park's Full Accuracy Design For Research On Stereotypes

Judd and Park (1993) developed the first componential model focusing on explaining sources of accuracy and inaccuracy in social stereotypes. Their approach was primarily designed to analyze accuracy and bias in at least two groups' perceptions of one another's attributes. Thus, their analysis incorporates four main components of judgments regarding groups: elevation, perceiver group, target group, and attributes and all possible two-way and three-way combinations of these components:

$$\begin{aligned} \text{Judgment} = & \text{constant} + \text{perceiver group effect(pge)} + \\ & \text{target group effect(tge)} + \\ & \text{attribute effect(ae)} + \\ & (\text{pge} \times \text{tge}) + (\text{pge} \times \text{ae}) + (\text{tge} \times \text{ae}) + \\ & (\text{pge} \times \text{tge} \times \text{ae}). \end{aligned} \quad (4)$$

Because Judd and Park's (1993) componential model is not identical to those of Cronbach (1955) and Kenny (1994), they are discussed here. Because they are so similar, however, and because there are so many more terms than in Cronbach's or Kenny's approaches, only a brief simplification of their main ideas is presented.

The constant is, once again, the mean level of judgments. The perceiver group effect refers to mean differences in the judgments of different groups of perceivers, the target group effect refers to the mean difference in how different groups of targets are judged, and the attribute effect refers to mean differences between attribute classes. The various two- and three-way interaction terms represent unique combinations of perceiver group, target group, and attribute, over and above the constant and the three main effects.

These components are then compared to groups' scores on the criterion. If one group of perceivers generally tends to overestimate targets' score on the criterion, there will be some perceiver elevation inaccuracy. If one group of targets is consistently overestimated, there will be some target elevation inaccuracy. If one class of attributes is consistently overestimated, there will be an attribute elevation inaccuracy. And so on.

Do people see other groups (in general) in a stereotypical manner? The attribute elevation effect represents an overall tendency to over- or underestimate the prevalence of a particular type or class of attributes across the target groups. If attributes are chosen such that they are stereotypic for one group and counterstereotypic for the other, the attribute elevation effect becomes a "stereotypicality" effect—the tendency to view groups as more or less stereotypic than they appear on the criterion. A general tendency to overestimate stereotypical attributes and underestimate counterstereotypical

attributes represents a general tendency (across target groups) for the stereotype to exaggerate real differences. A general tendency to underestimate stereotypical attributes and overestimate counterstereotypical attributes represents a general tendency (across target groups) for the stereotype to underestimate real differences.

For example, let's say perceivers are asked to rate men and women on the traits aggressive, strong, and mathematical. High attribute elevation means that people perceive a greater difference between men and women, across these traits, than really exists. This approach can also be used to assess valence inaccuracies (is a group seen more positively or negatively than they deserve?) and dispersion inaccuracies (is a group seen as more or less variable than they really are?).

Like the other componential approaches, Judd and Park's (1993) full accuracy design was modeled after an analysis of variance—but with three ANOVA factors (perceiver group, target group, attributes and all two-way and three-way interactions) rather than the two of Cronbach and Kenny's SRM. Although a full discussion of those factors is beyond the scope of this chapter, the three-way combination is particularly important to the study of stereotype inaccuracy because it tests for ingroup bias. The Subject Group \times Target Group \times Attribute factor tests whether stereotype exaggeration or underestimation (the attribute effect) is more likely to occur when people from group A judge people from group B and when people from group B judge people from group A than when people from group A judge people from group A and people from group B judge people from group B. If, for example, stereotype exaggeration only occurs when people judge groups other than their own, one would have an ingroup bias effect.

4. Components as a Means to Measure and Eliminate Noise In Judgments

Componential approaches are particularly useful for assessing and eliminating noise (judgment factors irrelevant to accuracy) from judgments. For example, consider a case in which both judgment and criteria are measured on a subjective scale, as might be the case when judging personality traits and the criteria are target self-reports. There could be a discrepancy between judgment and criteria not because perceivers are inaccurate but because they use the judgment scale in a different manner than do targets. Such subjective differences (or response biases) can be measured and eliminated when using a componential approach. Typically they manifest as one or more of the various elevation components in the different models (see, e.g., Kenny, 1994,

for a detailed discussion of how componential approaches can assess and remove response biases).

5. *Must Components Be Assessed In All Accuracy Research?*

Ever since Cronbach's (1955) review, researchers have emphasized the importance of assessing components, sometimes going as far as to claim or convey that components must be assessed to address accuracy questions (see, e.g., Hastie & Rasinski, 1988; Jones, 1990; Judd & Park, 1993; Kenny & Albright, 1987). Whether all accuracy research must assess components, however, depends on the precise meaning of this claim. Next, therefore, moderate and strong forms of this claim are considered.

a. The Moderate Form: Should Researchers Understand Existing Componential Approaches? The moderate form of this claim suggests that accuracy researchers should understand existing componential approaches to have better insights into the meaning of the results obtained from studies that do not explicitly assess components. This is undoubtedly true. It certainly behooves researchers interested in accuracy research to have more, rather than fewer, insights into the potential sources of social perception and the processes leading to accurate or inaccurate judgments and, especially, of the limitations and potential artifacts that influence whatever index of accuracy they do use.

b. The Strong Form: Process Versus Accuracy (Again). The strong form of the claim is that all accuracy researchers must explicitly perform componential analyses because otherwise their research will be meaningless or uninterpretable. A variety of logical, theoretical, and empirical considerations indicate that this strong claim is overstated.

First, componential approaches provide one class of explanations for how a person arrived at an accurate or inaccurate judgment. Indeed, Cronbach (1955) titled his article "*Processes Affecting Scores on 'Understanding of Others' and 'Assumed Similarity'*" (emphasis mine). Why? Because components provide information about the processes of judgment.

How do I arrive at my prediction that Mike Piazza will hit a grand slam? Do I always predict that he will hit home runs, or am I a particularly astute judge of Mike's hitting? Why does Louise think African Americans are less likely to complete high school than they really are? Does she underestimate every group's likelihood of completing high school? Does everyone, including African Americans, underestimate African Americans' likelihood of completing high school? Or is Louise ethnocentric, underestimating only African Americans' success and not her own group's success at completing high school? Componential analyses provide answers to these types of questions

by addressing the processes by which people arrive at accurate or inaccurate social judgments.

However, process is irrelevant with respect to establishing the degree of (in)accuracy of some perception. If I say "Mike is going to hit a home run" and he does, this particular prediction is right. End of discussion regarding my degree of accuracy. Although a single judgment is rarely psychologically interesting, the principle is just as true for multiple judgments. If I say that Mike Piazza is the best home-run-hitting catcher of the modern era and that he hits more home runs than any other catcher, again, I am right.

With respect to understanding how I arrived at that prediction, it would be valuable to estimate my elevation, stereotype accuracy, differential elevation, and differential accuracy (if you like Cronbach); or, if you prefer SRM, my elevation, my perceiver effect, Mike's target effect, and our interaction effect. But if one wants to determine whether my prediction is accurate, the only thing we need to do is figure out whether he hit the ball over the outfield fence in fair territory (in the case of the single judgment), and whether he hit more home runs during his career than any other catcher (in the case of multiple judgments).

If non-Armenians commonly believe that Armenians are public parasites burdening the financial community with their constant need for charity far more than other groups, and Armenians actually make fewer demands on public charity than other groups (LaPiere, 1936), then non-Armenians overestimate the financial burden created by Armenians. Again, period, the end, with respect to establishing the inaccuracy of this belief.

Interpreting that inaccuracy is another matter. As Ryan, Park, and Judd (1996) have pointed out, in the absence of their full accuracy componential design, we cannot conclude, as did LaPiere (1936), that this means those holding nasty beliefs about Armenians are necessarily anti-Armenian bigots. Perhaps they overestimate every group's need for charity, including that of their own. In that case, they are not ethnocentric at all. People with nasty beliefs about all groups including their own may be mean and evil, but they are not ethnocentric. Judd and Park's (1993) full accuracy design would be extremely useful for providing some insight into why they overestimate Armenians' request for charity, but it is completely irrelevant with respect to establishing whether anyone overestimates Armenians' requests for charity. That question can only be answered by comparing estimates of Armenian need for charity to some criteria.

c. No Reification of Components. It is extremely tempting to reify componential approaches to accuracy. First, they are statistically rigorous and complex, which gives them a particularly scientific aura. Second, they capture important, fundamental aspects of social perceptual judgment processes. Third, they successfully identify sources of bias or noise in judgments

that few researchers usually mean by accuracy. Thus, it is tempting to view components as concrete fixtures on the social perceptual landscape. If they are there, then should not accuracy researchers have to assess them?

Such absolutist positions regarding components ("Cronbach's [or SRM] components must always be assessed" or "accuracy can only be viewed componentially," etc.) are overstated for several reasons. First, there is no one right way to divide up components of social perception, as should be clear from this brief review of Cronbach's, Kenny's, and Judd and Park's componential approaches. They have important similarities, but, obviously, there are also differences between all three. Such differences are made salient when the three approaches appear side-by-side, as they do in Table II. If there were any single "right" set of components that "must" be examined, if components were actually hard and fast fixtures in the social perception landscape, there could not possibly be three different breakdowns of components unless one breakdown is "right" and the other two are "wrong" or unless each was woefully incomplete.

If all were partially right, but incomplete in that they failed to address components identified by other researchers, then a full componential model would need to assess all the components identified by all models. Such a model is presented at the bottom of Table II. If components are "real" and "must" be assessed, then the only complete way to do it would be to assess the more than 50 components identified in this model. Such a model has never been recommended even by advocates of componential approaches and is not being recommended here. Indeed, it is so extreme as to border on absurd—but such an absurd model might be required if all components "must" be assessed.

The situation, however, is far more complex than even this hypothetical combined componential model suggests. There is a potentially infinite number of ways in which social perception could be broken down into components (see also Kruglanski, 1989). Attributes could be further broken down into a variety of types or subclasses (e.g., positive vs. negative; explanations vs. descriptions vs. predictions; behaviors vs. traits, and so on). Similarly, both perceiver and target groups could be broken down not only by ingroup and outgroup but by any of the infinite ways of identifying groups (culture, demographic characteristics, memberships in organizations, professional expertise, etc.).

This is not meant to suggest, however, that existing componential approaches are purely subjective and arbitrary and, therefore, can be ignored, but the choice of components will depend entirely on the types of processes one would be interested in studying and the types of response biases one would like to assess or eliminate. Different componential breakdowns serve

TABLE II
COMPONENTIAL APPROACHES TO SOCIAL JUDGMENT

Cronbach's (1955) Components:

$$\begin{aligned} \text{Judgment of a Person's Trait} = & \text{Constant (elevation)} + \\ & \text{Target main effect (differential elevation)} + \\ & \text{Trait main effect (stereotype accuracy)} + \text{Target} \times \\ & \text{Trait interaction (differential accuracy)}. \end{aligned}$$

Social Relations Model Components (e.g., Kenny, 1994):

$$\begin{aligned} \text{Judgment of a Person's Trait} = & \text{Constant (elevation)} + \\ & \text{Perceiver main effect (perceiver effect)} + \\ & \text{Target main effect (target effect)} + \text{Perceiver} \times \\ & \text{Target interaction (relationship effect)}. \end{aligned}$$

Judd and Park's (1993) Components for Research on Stereotype Accuracy:

$$\begin{aligned} \text{Judgment of a Group's Traits} = & \text{Constant} + \text{Perceiver Group effect (pge)} + \\ & \text{Target Group effect (tge)} + \\ & \text{Attribute (stereotypic vs. counter-stereotypic) effect (ae)} + \\ & (\text{pge} \times \text{tge}) + (\text{pge} \times \text{ae}) + (\text{tge} \times \text{ae}) + \\ & (\text{pge} \times \text{tge} \times \text{ae}). \end{aligned}$$

Hypothetical Componential Approach Combining Cronbach's, Kenny's, and Judd and Park's Approaches:

$$\begin{aligned} \text{Judgement of a Person's Trait} = & \text{Constant} + \text{Perceiver Group effect} + \text{Target Group effect} + \\ & \text{attribute (stereotypic vs. counter-stereotypic) main effect} + \\ & \text{Target main effect} + \text{Perceiver main effect} + \\ & \text{individual trait effect} + 15 \text{ two-way interactions} + \\ & \text{all 20 three-way interactions} + \text{all 11 four-way interactions} + \\ & \text{all 4 five-way interactions} + \text{the six-way interaction}^a \end{aligned}$$

^aThis model assumes all main effect terms are independent, which may not be the case. For example, the target group main effect may not be independent of the target main effect and the attribute main effect may not be independent of the individual trait effect. In such a situation, there might be fewer interactions than displayed in this model. There would, however, still be literally dozens of such interactions. No researcher has ever advocated this model, including me.

different purposes and provide insights into different aspects and processes of social perception. Thus, understanding existing componential approaches would seem crucial to anyone studying accuracy to gain insights into how best to interpret their own or anyone else's data addressing the degree of (in)accuracy in social perception.

d. Componential Models May Be Most Important When the Criteria are Self-Reports and Self-Perceptions. Although Cronbach's componential approach never generated much empirical research, Kenny's and Judd and Park's have, and much of that research has used target individuals' or target groups' self-perceptions as criteria (see, e.g., reviews by Kenny, 1994; Ryan, 1995). When the criteria are self-perceptions, especially self-perceptions regarding traits, attitudes, or dispositions, rather than behaviors or other objective characteristics, using a componential analysis is highly likely to be the best (clearest, most informative) method of analysis.

Self-perceptions of traits, for example, typically have no objective referent. How extraverted is someone who rates themselves "5" on a 1-7 scale with endpoints labeled "not at all" and "very"? It is hard to say because each choice is subjective, in that each rater imputes their own meaning to each scale point (e.g., Biernat, 1995). Such differences in subjective meanings cloud the assessment of accuracy. Componential approaches, however, are particularly good at identifying differences in subjective meaning and removing them from estimates of accuracy (this is often captured by the various elevation components). Thus, it is probably a good idea to use a componential approach, if possible, almost any time one uses self-perceptions as criteria.

B. NONCOMPONENTIAL APPROACHES TO ASSESSING ACCURACY

It has just been argued that componential processes are not necessary for the assessment of accuracy. This argument, however, does not rest solely on a critical evaluation of the claim that "all accuracy research must perform a complex componential analysis." Instead, much of the best evidence for the idea that *componential analyses are not strictly necessary comes from the many noncomponential approaches to the study of accuracy that have made important and enduring contributions to understanding social perception.* The next section, therefore, briefly reviews three of the most influential noncomponential approaches.

The term "noncomponential" here is potentially misleading because it unfortunately implies that one can completely ignore the components issue. Even "noncomponential" approaches can themselves be considered to assess subsets of components in the various componential models, as shall be made explicit in the discussion that follows. The term "noncomponential" is used, therefore, to refer to approaches that assess accuracy without an explicit and intentional assessment of components. When, from a componential standpoint, such approaches only assess a subset of components in one or more of the componential models, this is pointed out explicitly below. What

may be lost by not performing a full componential analysis is also explicitly discussed.

1. Correlational Approaches

Most noncomponential approaches to assessing accuracy or processes underlying accurate and inaccurate social perceptions use correlations (typically Pearson's, occasionally regression coefficients) to assess the extent to which judgments correspond to criteria. In general, when judgments concern a single attribute, correlations between judgments and criteria capture Cronbach's (1955, p. 191) differential accuracy correlation, which he described as "sensitivity to individual differences. . . . These are the only processes included in present measures of social perception which depend on J's [perceivers'] sensitivity to the particular O [target]."⁵

The simplest and most typical form of correlation in accuracy research is that between a set of perceivers' judgments or predictions regarding a single trait or attribute of a set of targets and a criterion measure of that attribute or trait. For example, teachers' predictions regarding students' achievement may be compared with grades or standardized test scores; interviewers may evaluate a set of interviewees' skills, which are then compared with work samples; or perceivers may estimate the percentage of people belonging to various demographic groups that complete college, which are then compared to Census figures. Such correlations automatically remove the elevation and stereotype accuracy components from correspondence between perceivers' judgments and the criterion. (This is because these correlations reduce data to deviations from the mean.) Thus, a simple correlation (between judgment and criterion) goes a long way toward eliminating many of the biases, artifacts, and problems in assessing accuracy first identified by Cronbach.

Of course, the correlation coefficient is not perfect. First, it removes or avoids but does not directly assess Cronbach's (1955) elevation and stereotype accuracy components. Because correlations remove average differences between judgments and criterion, they cannot assess any consistent tendency to over- or underestimate targets (elevation, à la Cronbach). If it were important to assess those components to address some research question, *one could not use the correlation to do so.*

⁵Cronbach was, in addition to differential accuracy correlation, referring to differential elevation correlation in this quotation, which refers to the correlation between the perceivers' judgment of a target averaging over all attributes and the targets' average score on the criteria representing those attributes. This, however, is irrelevant when there is only a single attribute being judged.

Second, correlations equate the variability of judgments and criterion. Therefore, they cannot assess whether perceivers consistently over- or underestimate target variability.

Because mean and variability differences between judgments and criteria probably often reflect response bias or scaling discrepancies between perceiver and criterion, these limitations to correlations do not greatly undermine their utility in assessing accuracy. The term "scaling discrepancies" is used here to refer to the idea that people may use scale points in a manner differently than used in the criterion. This would obviously be true if, for example, judgment and criterion are assessed in different metrics (e.g., subjective rating scale and percentages, respectively).

People, however, still might use the numbers in some scale differently than is used for the criterion, even if they are supposedly on the same scale. For example, let's say Alfred estimates three people's IQ scores as 40, 50, and 60, when they are really 115, 120, and 125. Although it is possible that Alfred believes all three of these fairly intelligent people are neurologically impaired, it is more likely that Alfred does not fully understand how IQ scores are scaled. He dramatically underestimates people's IQ in absolute terms, but his estimates are also overly sensitive to actual variations in IQ (Alfred's judgments go up 10 IQ points for every 5-point increase in actual IQ). Given his subjective IQ scale, however, the correlation between his judgments and actual IQ would be perfect (1.0), because mean differences in judgment criteria are irrelevant to computation of the correlation, the correlation coefficient is computed after statistically equating the variability in judgment and criterion, and his judgments move in (differently scaled) lockstep with targets' actual IQ.

Thus, the correlation coefficient would yield a conclusion that Alfred is an excellent judge of people's intelligence. Is the conclusion justified? As long as one keeps in mind that what this really means is that "Alfred is very good at detecting differences in people's intelligence, but this does not tell us anything about either how Alfred uses the IQ scale or about whether he consistently over- or underestimates people's intelligence," it is perfectly justified.⁶

2. Construct Validity and Correlational Approaches To Accuracy

The section on criteria argued that assessing accuracy was much like assessing the validity of many social science constructs. This is also important here, because the correlation coefficient is so frequently used to

⁶In empirical studies, one also needs to be concerned with the level of analysis and nonindependence issues that arise when, for example, each perceiver judges several targets. This, however, is a design and statistical issue that is beyond the scope of this chapter (but see Kenny, Kashy, & Bolger, 1998, for a detailed discussion of this issue).

establish the validity of some measure that it is often referred to as the "validity" or "validity coefficient" of some measure (e.g., Campbell & Stanley, 1963; Cook & Campbell, 1979; Dawes, 1979). In much the same manner, correlations can be used to assess the accuracy of social perception.

Establishing accuracy, however, is somewhat more complex than establishing construct validity because social perception, judgment, and expectations are themselves constructs. Thus, all the issues involved in establishing construct validity arise not just when measuring targets' attributes but when measuring perceivers' expectations, perceptions, judgments, and beliefs about others. Accuracy, therefore, is not usually best reflected by correlations between observable measures (e.g., a measure of perceiver expectations and a measure of target extravertedness). Accuracy will often best be reflected by correlations between the underlying constructs representing the social perception and the criteria (this can often be accomplished either by disattenuating correlations for unreliability or by using LISREL-type models; see, e.g., Bollen, 1989; Carmines & Zeller, 1979).

This should not, however, be misinterpreted to mean that all accuracy research must necessarily assess relationships between underlying constructs rather than observed measures. Sometimes it may just not be possible to do so, as, for example, when there is only a single item assessing either the perception or the criterion. This, of course, does not constitute any sort of immovable obstacle to or fatal flaw in accuracy research. Correlations among such measures, however, will tend to underestimate accuracy to the extent that the observed measures only imperfectly reflect the underlying attributes or expectations. Therefore, people may be more accurate than indicated by research that only assesses correlations between observed measures of expectations and observed measures of criteria.

This brief delving into construct validity, correlations, and unmeasured variables was necessary to lay the foundation for understanding three of the main noncomponential approaches for assessing accuracy in social perception. All three are fundamentally based on the correlation between perception and criteria.

3. Brunswik's Lens Model

Brunswik (1952) suggested that accurate perception of reality (both object and social) involves the use of cues probabilistically related to objective reality. He metaphorically called his approach the Lens Model to capture the idea that objective reality is never observed directly. Instead, cues related to objective reality must be observed and interpreted as relevant to some judgment; that is, objective cues are seen through the "lens" of subjective perception and judgment.

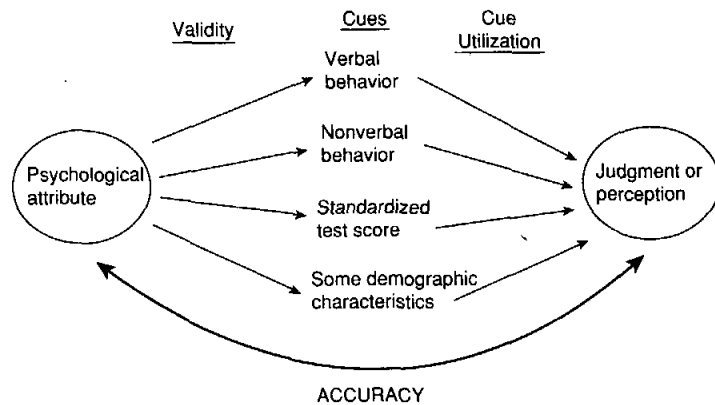


Fig. 1. Brunswik's Lens Model. Single-headed arrows represent causes. The double-headed arrow represents the correlation. This figure depicts a general example of how the Lens Model might be used. It can be used for assessment of accuracy in perception of physical attributes (e.g., size, distance) as well as of psychological attributes. Furthermore, the four cues shown here are simple concrete examples of potential manifestations of some attribute; the Lens Model is not restricted to these types of cues. In Lens Model research, what is called accuracy here is usually called "achievement."

This does not necessarily mean that perception is a purely subjective phenomenon unrelated to objective reality. Indeed, one of the main purposes of the Lens Model is to provide a mechanism not only for assessing people's degree of accuracy but also for understanding sources of both accuracy and inaccuracy in their judgments. The Lens Model has been used to study accuracy, inaccuracy, and bias in a wide variety of business, educational, criminal, and interpersonal contexts (e.g., Bernieri & Gillis, 2001; Craik & Appleyard, 1980; Reynolds & Gifford, 2001; Schmitt, Noe, & Gottschalk, 1986).

Figure 1 presents a simplified but general version of the Lens Model. It captures several main ideas. First, the circled "Psychological Attribute" represents some sort of psychological construct that cannot be directly observed (self-esteem, extraversion, intelligence, etc.). The Cues, shown in the middle of the figure, are directly observable or measurable phenomena. The arrows pointing from the Psychological Attribute to the Cues are labeled "Validity," because they represent the extent to which the underlying attribute manifests itself in the observable Cues.

The rightmost circle represents perceivers' judgments (or perceptions). The arrows going from the Cues to Judgments represent the extent to which the observable cues influence perceivers' judgments (labeled Cue Utilization). The Lens Model characterizes social perception as a two-step process: observable manifestation of psychological attributes and perceiver use of

observable cues to arrive at judgments. Accuracy, therefore, is captured by the correlation between the psychological attribute and the judgments (the long, double-headed arrow in Fig. 1). Correlations assess how well the judgments correspond to the attributes—that is, accuracy.

The Lens Model is a noncomponential, correlational model for assessing both degree of accuracy and processes of social perception. Identifying cue validity and cue utilization focuses on a very different set of the processes than is typically the focus in componential models. As such, it provides different (not better or worse) types of insights into processes of social perception than do componential models.

4. The Realistic Accuracy Model

Funder's (1995, 1999) Realistic Accuracy Model (RAM) draws on the set of fundamental assumptions described under "probabilistic realism" to create a model that could be viewed as an extension and elaboration of Brunswik's Lens Model. Some of the main ideas of RAM are depicted in Fig. 2. As with the Lens Model, overall accuracy is typically assessed by the correlation of the underlying attribute with the perceiver's judgment (represented by the large, curved, double arrow on the bottom of Fig. 2). Four steps needed for perceivers to arrive at an accurate judgment are displayed in between the underlying attribute and the judgment.

First, the underlying attribute needs to create some sort of observable evidence relevant to that attribute (the cues, in the Lens Model). Dishonesty, for example, is not likely to be displayed in a large lecture hall (except perhaps during test time). Interest in the class is more likely than honesty

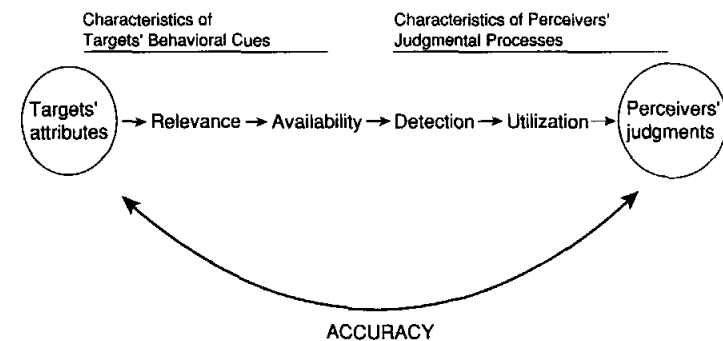


Fig. 2. Funder's Realistic Accuracy Model. The single-headed arrows here represent steps in a sequence, not causal effects. The double-headed arrow represents the correlation between targets' attributes and perceivers' judgments.

to be displayed, for example, through high attendance levels, keeping up with class assignments, or class participation.

Second, the evidence needs to be at least hypothetically available to the perceiver. Whether or not a student has completed all assigned readings may rarely be directly observable to the class's teacher. Attendance and participation, however, would be considerably more available.

Third, the perceiver has to detect that evidence. In a large lecture hall, which students participate in is pretty obvious. However, detecting precisely which students do and do not attend regularly, out of a swarming mass of hundreds of students in the lecture hall, is considerably more difficult.

Fourth, the perceiver has to actually use the detected evidence/cues (and weight them appropriately) for arriving at an accurate judgment. If lecture hall teachers cannot remember which students regularly participate, it would be difficult to use participation as a basis for judgment regarding interest in the class.

Although this is the heart of Funder's (1995) RAM, applying the model might be considerably more complex. People have many underlying characteristics. Funder (1995) focused primarily on personality traits, but RAM seems applicable to all sorts of unobservable personal characteristics, including, for example, emotions, attitudes, motivation, and so on. Furthermore, one attribute may create many cues (as is most obviously captured in the Lens Model), and some cues may reflect multiple attributes. Thus, one type of complexity involves the sheer number of possible interrelationships among attributes, cues, and judgments.

Like several of the componential models, RAM also considers how the perceiver, target, attribute, and evidence relate to accuracy (Funder [1995] referred to these as Judge, Target, Trait, and Information, respectively, but I am sticking with the terminology I have used throughout this chapter). This is not to identify components, however. Instead, the purpose is to analytically identify how specific combinations of perceiver, target, attribute, and evidence might combine to influence accuracy.

For example, some perceivers may be particularly good (poor) at evaluating certain types of traits (e.g., clinical psychologists might be better than most of the rest of us at evaluating others' mental health). Some perceivers might be particularly good (poor) at judging particular targets (e.g., close friends might be better judges of each other than are strangers). Some perceivers might be especially good (poor) at using or obtaining certain types of evidence (e.g., some people may be better than others at picking up on targets' emotion-revealing nonverbal cues). Funder (1995, 1999) has discussed at considerable length how the various combinations of perceiver, target, attributes, and evidence may combine to influence accuracy.

Like the Lens Model, RAM assumes that relationships between underlying attributes, cues, and judgments are probabilistic. Like the Lens Model,

overall accuracy is typically assessed via correlations, although discrepancy scores (between judgment and criterion) can be used, too (see Funder, 1987, 1995). RAM is particularly good at explaining why accuracy in person perception may often be low. For the judgment to closely correspond to the criterion, that criterion needs to clearly manifest itself in ways that could be and, in fact, are detected by the perceiver, and then the perceiver needs to use that detected information (as well as not use information that is not relevant to the judgment). A breakdown at any step will dramatically undermine accuracy. Furthermore, by focusing on combinations of perceiver, target, attributes, and evidence, RAM is also particularly good at highlighting processes that may enhance or undermine accuracy.

5. Dawes's Improper Linear Models

Dawes (1979) made a very interesting discovery. In reviewing his own and others' research on decision-making, he found that people tend to be very good at identifying the evidence or cues that are relevant to making some prediction but are not very good at combining or integrating those cues. Thus, their overall predictive accuracy tends to be quite low. Note, however, that this is not because people are completely inept. They are good at one part of the prediction task (identifying criteria for making a prediction) but poor at another part (putting those criteria together).

One of Dawes's (1979) examples involved admissions to graduate school in psychology. The criteria typically used for making admissions decisions seem appropriate: GRE scores (general intellectual ability), GPA (achievement at academic tasks over an extended period), and letters of recommendation (what experts in the field who are highly familiar with the applicant have to say about him or her). Nonetheless, Dawes (1979) found that the correlation of graduate admissions committee evaluations with later success in graduate school is typically quite low (.19).

If people were completely inept, they would not even use appropriate criteria; that is, the criteria they do use would not predict success in graduate school. However, if they are good at identifying the appropriate criteria, but use them poorly, then the raw criteria themselves should do a much better job at predicting graduate success. This was indeed the case—the overall (multiple) correlation of the criteria themselves with graduate success was about .4.

What to do? It is unreasonable to expect admissions committees to compute complex statistical formulas in their heads or to create a formal statistical score for each applicant. Dawes provided an elegantly simple and even amusing answer. Identify the criteria, weight them all equally, and add. For example, GRE, GPA, and letters of recommendation might each be transformed into standardized scores (obviously, letters would need to first be placed on some sort of rating scale). Priscilla, with good GREs, a high

GPA, and excellent letters of recommendation might receive *z*-scores of 1, 2, and 2, respectively, for a total score of 5. George, with high GREs, a good GPA, and good letters, might receive *z*-scores of 2, 1, and 1, for a total of 4. Priscilla would be ranked higher than George.

This is different from a formal statistical model primarily because the weights for each predictor have been chosen in a less than optimal manner (many statistical prediction techniques, such as regression, identify how to weight the criteria in such a manner as to maximize their overall predictive validity). Dawes's (1979) analysis showed, however, that this "suboptimal" weighting of criteria predicts outcomes nearly as well as do formal statistical models. In the graduate admissions example, Dawes' improper linear model correlated .38 with future success in graduate school. Dawes (1979) went on to show that a simple, improper linear model performed similarly well in predicting all sorts of outcomes, including choice of bullet-type for a police department and a bank's predictions regarding companies likely to go bankrupt.⁷

Dawes' improper linear model is fundamentally different than the Lens Model and RAM. The Lens Model and RAM were specifically designed to assess degree of accuracy and processes underlying social perception. That is, they were meant to describe aspects of the social perception process. In contrast, Dawes' model is primarily prescriptive (it suggests how people should go about making decisions and arriving at predictions).

Nonetheless, it has been included here for two reasons. First, Dawes's (1979) conclusion that people are good at selecting criteria, but not good at using them, is descriptive. In RAM terms, it indicates that people often are good at detecting available and relevant cues, but that they often do not use them well (in Lens Model terms, their cue utilization would be poor). Second, although Dawes did document that people were, on their own, not very good at arriving at accurate predictions, he also showed that the accuracy of their predictions could easily be improved. Identify the criteria, weight them equally, then add!

C. COMPONENTIAL AND NONCOMPONENTIAL APPROACHES TO THE STUDY OF ACCURACY

Componential approaches to the study of accuracy have provided several major contributions to the study of social perception. As a research tool for assessing the processes of social perception, they demonstrate again how

⁷Dawes's results were actually even more striking. The equal weight improper linear model generally (i.e., not just in the graduate admissions example) predicted outcomes better than the split sample cross-validated regression weights.

accuracy research typically provides the type of double contribution that few, if any, experimental studies focusing exclusively on error, bias, and process can provide. Componential approaches address accuracy, inaccuracy, and process. They are particularly useful at identifying some of the reasons a perceiver is a particularly good or poor social judge and also can demonstrate that some perceiver might be accurate in some ways, yet inaccurate in others, even when judging the same target. Last, they are especially good at measuring and eliminating response biases that are largely or completely irrelevant to the assessment of accuracy.

It is also clear, however, that componential approaches have not been the only way to productively study accuracy. Three noncomponential approaches, Brunswik's (1952) Lens Model, Funder's (1995) RAM, and Dawes's (1979) improper linear models were also discussed. All rely on the correlation as the primary means for establishing correspondence between social judgment and criteria, and all have provided important insights into both the (in)accuracy of social perception and basic social perceptual process. Other approaches relying on Bayes's theorem (e.g., Krueger, 1996; Krueger et al., 2003; McCauley & Stitt, 1978; McCauley et al., 1980) and structural equation modeling (e.g., Jussim, 1991; West & Anderson, 1976; Williams, 1976) have also contributed to understanding processes leading to accurate and inaccurate social perceptions, especially with regard to interpersonal expectations and social stereotypes.

It is clear, therefore, that although componential approaches are valuable and important, one need not perform any particular componential analysis to assess accuracy. Indeed, in the absence of data permitting a full componential analysis, performing a simple correlation between belief and criterion, though imperfect, would typically provide far more information than would doing nothing. Even in such a situation, however, it behooves the researcher performing the correlation to understand componential approaches to appropriately interpret just what that correlation does and does not show.

IX. Conclusions

This review has documented many of the conceptual, methodological, and statistical advances in the psychological study of social perceptual accuracy, with a particular focus on addressing many of the criticisms of accuracy research that have been raised over the last 50 years. In doing so, the conclusion has been repeatedly reached that, although there are genuine difficulties, ambiguities, or obstacles to accuracy research, they are no greater than in most other areas of social psychological research. Indeed, this

review has shown that the conceptual, theoretical, methodological, and empirical bases of accuracy research are identical to those of much other research in psychology.

By taking stock of the many conceptual and methodological advances in accuracy research that have been made over the last 20 years, this review has attempted to provide a road map for further accuracy research. Although it is impossible for such a map to represent every twist and turn in the road, it has attempted to provide some valuable signposts providing direction in the pursuit of accuracy research. These signposts show how to circumvent or avoid many known obstacles, indicate that the road is clear in many directions that were once believed to be filled with obstacles, and have identified many alternative routes available when known obstacles arise.

A. NONE OF THE COMMON CRITICISMS OF ACCURACY RESEARCH THREATENS ITS VIABILITY

This review has contested the viability of some of the common theoretical, conceptual, and methodological criticisms of accuracy research. Some criticisms were seen to be manifestly false; others were seen to be logically incoherent. Yet others were seen to be partially valid but not to constitute serious threats to accuracy research. This review also showed that even completely valid criticisms of accuracy research could readily be addressed by a clear understanding of the nature of, and limitations to, the insights into social perception that accuracy research can provide.

B. THE CRITERIA FOR ACCURACY ARE IDENTICAL TO CRITERIA USED IN PSYCHOLOGICAL, SOCIAL, AND BEHAVIORAL SCIENCE RESEARCH

The crucial issue of identifying criteria for establishing accuracy was also explored. This review contested arguments criticizing accuracy research for alleged difficulties in identifying appropriate criteria for validating social beliefs. Such criticism would seem to logically require a similar criticism of nearly all research in psychology (with the possible exception of biological research). The logic of establishing accuracy of social perception overlaps almost completely with the (far less controversial) logic of establishing construct validity in the social sciences.

C. BOTH COMPONENTIAL AND NONCOMPONENTIAL APPROACHES TO ACCURACY ARE VALUABLE

Furthermore, several major approaches to the assessment of accuracy and processes leading to accurate and inaccurate social judgments were reviewed. This included reviews of the three main componential approaches to accuracy and three of the major noncomponential approaches to the study of accuracy. It was concluded that both componential and noncomponential approaches have made major and significant contributions to understanding social perception and accuracy.

D. PROCESS RESEARCH ON ERRORS, FLAWS, AND BIASES IS NOT MUTUALLY EXCLUSIVE WITH ACCURACY

Despite the often-demonstrated existence of a slew of logical flaws and systematic biases in lay judgment and social perception, such as the fundamental attribution error, false consensus, over-reliance on imperfect heuristics, self-serving biases, and so forth, people's perceptions of one another are surprisingly (though rarely perfectly) accurate. Despite being most well-known for their self-fulfilling effects, teacher expectations, for example, typically predict student achievement more because they are accurate than because of expectancy effects (e.g., Brophy & Good, 1974; Brophy, 1983; Eccles & Wigfield, 1985; Jussim, 1991; West & Anderson, 1976). People are also often surprisingly good at judging one another's personalities (e.g., Funder, 1987, 1999; Funder & West, 1993; Paunonen, 1991). People in close relationships are often better than strangers at inferring each other's thoughts and feelings (Ickes, 1993, 1997; Stinson & Ickes, 1992). Indeed, people even show some accuracy under very poor or minimal conditions in arriving at initial impressions among strangers (Kenny, 1994), for example, and in predictions based on very small samples of behavior (Ambady & Rosenthal, 1992).

E. BELIEFS ABOUT SOCIAL GROUPS (STEREOTYPES) OFTEN SHOW MODERATE TO HIGH ACCURACY

Perhaps most surprisingly, considerable research has accumulated showing that people's perceptions of groups and individual targets from differing race, sex, and social class backgrounds are often quite (though

not perfectly) accurate (Diekmann et al., 2002; Hall & Carter, 1999; Judd et al., 1995; Jussim et al., 1996; Lee et al., 1995; Krueger, 1996; Madon et al., 1998; McCauley & Stitt, 1978; Ryan, 1996, 2002; Ryan & Bogart, 2001; Swim, 1994; Williams, 1976). Because, however, accuracy is not mutually exclusive with, and sometimes is even enhanced by, bias, this claim does not refute or deny the existence of prejudice or discrimination. Indeed, many of the same studies demonstrating moderate to high accuracy have also demonstrated that stereotypes do sometimes lead to systematic biases in judgments of groups and individuals.

F. UNDERSTANDING ACCURACY IS ESSENTIAL FOR UNDERSTANDING AND ALLEVIATING SOME SOCIAL PROBLEMS

Research on accuracy would appear essential in the application of psychology to the remediation of many social problems, especially those involving intergroup relations. Beliefs about groups have for decades been presented as either a cause of social inequalities or as a reflection of prejudice and existing dominance hierarchies, or both (e.g., Allport, 1958; American Psychological Association, 1991; Fiske & Taylor, 1991; Jost & Banaji, 1994; LaPiere, 1936; Pickering, 2001; Sidanius & Pratto, 1999). If such beliefs are largely erroneous, and if they cause injustices, as many such perspectives imply, then the solution to many group-based social problems should focus almost entirely on correcting the beliefs and perceptions of those holding such beliefs. If the problem is erroneous perception, then correcting the perceptions should eliminate, or at least greatly alleviate, that problem.

To the extent, however, that some beliefs about group differences are accurate because groups really do differ on some societally valued attributes, then correcting erroneous perceptions would be a misguided endeavor, or at best a woefully incomplete solution. In this situation, efforts need to be focused on redressing the underlying inequalities, which will probably involve devoting resources to improving the station of lower status or unjustifiably stigmatized groups. Accuracy research, therefore, would appear to be crucial with respect to identifying how to devote resources to the alleviation of social problems—with relatively more resources being directed to correcting perceiver beliefs when those are shown to be inaccurate but relatively more resources being directed to improving the station of low status groups when perceiver beliefs are shown to be accurate.

G. AN INVITATION TO STUDY ACCURACY

There has been a great blossoming of accuracy research over the last 20 years (Ambady & Rosenthal, 1992; Funder, 1987, 1995, 1999; Ickes, 1993, 1997; Kenny, 1994; Kenny & Albright, 1987; Judd & Park, 1993; Jussim, 1989, 1991; Jussim et al., 1996; Lee et al., 1995; Oakes et al., 1994; Ryan, 2002; Swim, 1994). Now that the major obstacles to accuracy research have been cleared away and that new and innovative ways to address accuracy are being continuously developed, if anything, the pace of research on accuracy should accelerate over the next 20 years.

Nonetheless, specific future developments are difficult to anticipate. Twenty years ago, I would never have predicted that people could be as accurate as they apparently are on the basis of minimal information (Ambady & Rosenthal, 1992) or that stereotypes would show as much accuracy as they consistently seem to show (Brodtt & Ross, 1998; Hall & Carter, 1999; Lee et al., 1995; McCauley & Stitt, 1978; McCauley, 1995; Ryan, 2002; Swim, 1994; Wolsko et al., 2000). Nonetheless, many classic issues and recent developments in social psychology naturally raise a host of interesting and important questions about accuracy.

1. Moderators: The Person, the Situation, and the Relationship

Social and personality psychology have a long tradition of examining the role of individual differences, situational forces, or both in causing or moderating all sorts of social phenomena. This tradition naturally lends itself to the study of accuracy. Hall and Carter (1999), for example, examined whether a slew of individual difference factors increased or reduced the accuracy of people's beliefs about differences between men and women (the most successful moderator was nonverbal sensitivity). With the modern methods that have cleared away the old obstacles, we may finally be able to obtain some answers to the question that dominated early accuracy research and seemed to have no answer: "Who is a more accurate judge of others?"

Similarly, researchers have also begun addressing the other classic social psychological issue—situational moderators of accuracy. Wolsko et al. (1999), for example, found that White students, given a "multicultural" set (one emphasizing the importance of understanding and valuing bona fide racial and cultural differences), held more accurate racial stereotypes than those given a "color-blind" set (one emphasizing the sameness of all people and deemphasizing the importance of group differences). Those given

the multicultural set also used their stereotypes more appropriately when judging individual targets. Thus, situations evoking different sociopolitical mindsets moderated stereotype accuracy.

Others areas ripe for an investigation of situational moderators of accuracy include interpersonal expectations in general and teacher expectations in particular, managerial perceptions of employees, and first impressions.

Accuracy may also vary by aspects of relationships. Length of relationship seems like an obvious moderator (the longer people get to know one another, the more accurately they should see one another). Although research on stereotypes is broadly consistent with this hypothesis (especially the overwhelming evidence that people rely less on stereotypes the more individuating information they have; Eagly, Makhijani, Ashmore, & Longo, 1991; Kunda & Thagard, 1996), research using the SRM has not always found that accuracy increases with length of relationship (Kenny, 1994). Whether accuracy also systematically varies by type of relationship (e.g., romantic, parent-child, occupational, etc.) is currently unknown.

2. Accuracy and Bias and Their Integration

Whether lay social beliefs and perceptions are typically rational and accurate or inaccurate and biased has been controversial for decades. Clearly, however, those emphasizing error, bias, and the ways in which social beliefs create social reality have dominated the literature on social cognition (e.g., Fiske, 1998; Jones, 1986; Kahneman & Tversky, 1973; Nisbett & Ross, 1980; Snyder, 1984). These views have created an image of a social perceiver whose misbegotten beliefs and flawed processes construct not only illusions of social reality in the perceiver's own mind but actual social reality through processes such as self-fulfilling prophecies. In this bleak view, the mind becomes primarily a product of cognitive shortcomings and distorted social interactions.

Implicit in the prevalence of testaments to the power of error and bias is that something important has been discovered. Because error, bias, and accuracy can coexist simultaneously, however, merely demonstrating error or bias does not demonstrate inaccuracy. Although people undoubtedly commit errors and biases and are rarely perfectly accurate, almost none of the literature routinely cited as testaments to the power and prevalence of error and bias actually tests for accuracy. As a consequence, despite the manner in which it is cited, that literature provides little direct information about accuracy. One has to actually test for accuracy to reach conclusions about inaccuracy.

Putting at least some of the huge existing literature on error and bias into some context by also testing for accuracy may be less difficult than it seems,

because methods are being developed for extracting information about accuracy, agreement, and objectivity from studies that have focused exclusively on inaccuracy, disagreement, and subjectivity (Jussim, 2004; Jussim, Harber, Crawford, Cain, & Cohen, in press). Such methods suggest, for example, that classic demonstrations of bias and subjectivity, such as Rosenhan (1973) and Hastorf and Cantril (1954), actually provide far more evidence of accuracy and objectivity than of bias or subjectivity.

At least under some conditions, furthermore, heuristics, biases, and accuracy may be integrated. Many of the classic treatises on flaws and biases (e.g., Kahneman & Tversky, 1973; Nisbett & Ross, 1980) include disclaimers and phrases along the following lines: The many biases, flawed processes, and faulty heuristics we have identified may often serve people quite well. Although the error and bias classics rarely expanded much on this type of claim to show how, why, or when it might actually be correct, the claim itself is correct, for a variety of reasons documented here: bias will sometimes increase accuracy, a flawed or imperfect process may still be good enough to generate mostly accurate predictions or judgments, and expert models are sometimes contradictory, implying at least one is wrong or not applicable, so that demonstrating that people deviate from such models does not always mean they are inaccurate.

Both theory and empirical research have begun to address the role of heuristics and biases in increasing accuracy (Dawes, 1989; Jussim, 1991; Kenny & Acitelli, 2001). Indeed, it seems likely that the fact that many heuristics and biases do increase accuracy will help explain why they develop and are difficult to eliminate. As such, a deeper understanding of the social cognitive processes underlying judgment, prediction, and decision-making will undoubtedly be obtained by research seeking to understand the benefits, and not just the costs, of well-documented biases and heuristics.

3. Accuracy of Implicit Beliefs

The type of reaction time, categorization, and priming techniques common to research on implicit and automatic social cognition (Bargh & Chartrand, 1999; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998) at first glance may not seem to readily lend themselves to the study of accuracy. It is not meaningful, for example, to discuss the "accuracy" of a reaction time. Furthermore, much of this type of research has focused on attitudes, and attitudes cannot be accurate or inaccurate.

Nonetheless, research on the accuracy of implicit and automatic beliefs can and should be conducted. One key question involves the extent to which implicit and automatic beliefs correspond to social reality. For example, do

people's implicit beliefs about differences between groups correspond to the size of the actual group differences? To address this type of question, people's scores on the measures reflecting their implicit beliefs (priming effects, implicit associations, etc.) can be correlated with criteria regarding those beliefs. For example, implicit beliefs about differences between African Americans and Whites can be correlated with Census reports, or implicit beliefs about a particular person can be correlated with that person's attributes. Much of the work to date on implicit beliefs and attitudes emphasizes their dark side, in large part, because it has often focused on prejudice (Fazio et al., 1995; Greenwald et al., 1998). This emphasis is well-justified by the damage that prejudice wreaks on individuals and groups and is broadly consistent with social psychology's historical emphasis on error and bias. Nonetheless, the extent to which implicit beliefs about the self, other individuals, and groups reflect social reality (in addition to, and, perhaps, sometimes rather than) perceivers' errors and biases is currently unknown.

4. *Evaluating the Effectiveness of Interventions Designed to Increase Accuracy*

Accuracy interventions, although they are rarely referred to in those terms, pervade everyday life:

- **Psychotherapy:** The goals of personal insight, recognition of dysfunctional behavior patterns and ways to avoid them, and the reframing of unjustifiably damaging views of self or others are often goals of psychotherapy.
- **Education:** A liberal arts background, especially with experience in the humanities and social sciences, can be viewed as, in part, one immense undertaking intended to dissuade students of their inaccurate and parochial views of themselves, their peers, their country, and the various religions, cultures, nationalities, and ethnicities of the world. Advanced instruction in social, organizational, and cognitive psychology; business programs; statistics; and philosophy often provides specific training relevant to eliminating many of the errors and biases that are so well documented in the psychological literature.
- **Multiculturalism/diversity:** Interventions in both business and educational settings designed to reduce prejudice and intergroup conflict by increasing support for multiculturalism and diversity have become commonplace. Multiculturalism/diversity typically has a moral component ("we should value difference") that is beyond the scope of this chapter, but multiculturalism also has an accuracy component ("we should replace our false beliefs about other groups with more accurate ones"). This latter point is rarely put so bluntly, but how can we be "sensitive to others' cultural

differences" if we do not have some valid insights into and understandings of those differences? This type of point is frequently made in pleas for people to become more multiculturally sensitive. For example, Pinderhughes (1989) titled her book, which is essentially an extended treatise advocating multiculturalism in therapy, *Understanding Race, Ethnicity, and Power*. Presumably, she believes that such understanding can be attained, otherwise she would not have written the book or given it that title.

From the perspective here, this constitutes a call for an increase in the accuracy of people's beliefs about (understanding of, insight into) those from other cultural backgrounds. I would call that an increase in the accuracy of their stereotypes, because I simply define a stereotype as a belief about a group. Indeed, Pinderhughes (1989, p. 147) makes essentially the same points when she identifies several abilities that enhance people's multicultural competencies, two of which are "The ability to control, and even change false beliefs, assumptions, and stereotypes" and "The ability to respect and appreciate the values, beliefs, and practices of persons who are culturally different."

Respecting others' values, beliefs, and practices makes sense primarily if one has a reasonably clear (accurate) sense of what those values, beliefs, and practices are.

Unfortunately, however, the accuracy of the new beliefs created by interventions such as therapy, education, and multicultural programs is rarely tested. The validity and effectiveness of psychotherapy remains controversial (Dawes, 1994; Seligman, 1995). The specific issue of whether clients' beliefs typically become more accurate through therapy is largely unknown. As a consequence, this would appear to be an interesting and important area for future research.

Similarly, little research has evaluated the extent to which education improves accuracy. On the one hand, direct explicit statistical training increases people's ability to use appropriate statistical reasoning when solving problems (Fong & Nisbett, 1991). On the other, the overwhelming majority of studies demonstrating biased judgments, overreliance on faulty heuristics, and a myriad of failures to understand or apply principles of statistics or logic in everyday reasoning problems have relied on college student samples (e.g., Kahneman et al., 1982; Nisbett & Ross, 1980). Whether people who have received a college education generally reach more accurate conclusions about their friends, coworkers, and peers and about the larger groups in society than do their peers who have not gone to college but who are otherwise similar in age and intelligence is unknown.

Presumably, the thousands of diversity and multiculturalism workshops, classes, and programs that have been instituted on campuses around the

country do achieve some degree of reduction in discrimination, if only because they convey the message that "people in power around here want you all to get along." Furthermore, they are probably superior to doing nothing at all, but this is not guaranteed. Potential negative side effects, such as a dampening of the free exchange of views crucial to both a college campus and a functioning democracy, and an increase in the willingness to accept immoral or illegal behavior (e.g., acceptance of sexual harassment because someone is from a culture emphasizing machismo) "because that's how they do things in that culture," must be evaluated alongside any reductions in prejudice or discrimination. Nonetheless, the effectiveness of such programs has almost never been evaluated (e.g., McCauley, Wright, & Harris, 2000). Thus, their effectiveness in leading people to become more acutely aware of (i.e., to more accurately perceive) how the beliefs, values, and practices of people from various backgrounds differ from their own is both unknown and a potentially rich area for future research.

H. "THE COMMON AND MAINLY VERIDICAL CHARACTER OF THE BASIC HUMAN PERCEPTIONS"

This review has suggested that it is no longer justified to dismiss this blossoming of accuracy research for reasons such as accuracy research provides little information about psychological processes (process is a central focus of most accuracy research); research on accuracy is unnecessary because research on social cognition already shows that social perception is generally inaccurate (social cognition research has uncovered a host of errors and biases, but these do not demonstrate that people are generally inaccurate); research on accuracy often has not addressed how perceivers explain others' behaviors and attributes (it often has not, and this is a limit, but not a threat, to existing accuracy research); self-fulfilling prophecies may account for that which is "accurately" perceived (they may, but this does not preclude the study of accuracy); research on stereotype accuracy is unnecessary because even a stereotype that fits a group will inaccurately describe most members of that group (because this claim confounds levels of analysis, its challenge to stereotype accuracy research fails); it reifies individual and group differences (existing scholarship on accuracy has never claimed that accurately perceived differences constitute immutable characteristics); it is hopelessly politicized (it sometimes evokes hostile, politically motivated criticisms, and the complex interrelationships of accuracy, bias, and discrimination have received little discussion in the scholarly literature or even in everyday cultural discourse, but the research itself is probably no more

politicized than many other areas of social psychology); it exacerbates social problems (it is a powerful tool for the redress of many social problems); criteria are so fundamentally flawed or ambiguous as to undermine the viability of accuracy research (the criteria are identical to those used to assess self-fulfilling prophecies and to assess the validity of most psychological constructs); or much accuracy research has failed to perform a componential analysis (much has, and even when it has not, accuracy has been productively studied anyway).

Fifty years ago, Floyd Allport (1955), in responding to the New Look in Perception research which, arguably, launched the now-five-decade-long social psychological emphasis on error and bias, made the following appeal:

"What we are urging here is that social psychologists, in building their theories of perception, assume their share of the responsibility for reconciling and integrating their 'social-perceptual' concepts, fraught with all their deviations and special cognitive loadings, with the common and mainly veridical character of the basic human perceptions" (p. 372).

It now seems clear that the meta-theoretical, theoretical, and empirical paths have been sufficiently cleared of obstructions for accuracy research to provide the balance to psychological perspectives on social perception that Floyd Allport called for 50 years ago.

Acknowledgments

I sincerely thank Richard Ashmore, Gretchen Chapman, David Funder, Joachim Krueger, and Clark McCauley, for comments on earlier drafts of this chapter. This chapter is based, in part, on sections from a book by Lee Jussim, currently in progress and tentatively titled *Social Beliefs and Social Realities*.

References

- Allport, F. H. (1955). *Theories of perception and the concept of structure*. New York: Wiley.
- Allport, G. W. (1958). *The nature of prejudice* (2nd ed.). Garden City, NY: Doubleday.
- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*, 256-274.
- American Psychological Association. (1991). In the Supreme Court of the United States: *Price Waterhouse v. Ann B. Hopkins* (Amicus curiae brief). *American Psychologist*, *46*, 1061-1070.
- Asch, S. E., & Akert, R. M. (1977). Words and everything else: Verbal and nonverbal cues in social interpretation. *Journal of Personality and Social Psychology*, *35*, 443-449.

- Aronson, E. (1999). *The Social Animal* (8th ed.). New York: Worth Publishers.
- Asch, S. (1955). Opinions and social pressure. *Scientific American*, *11*, 31-55.
- Ashmore, R. D., & Del Boca, F. K. (1981). Conceptual approaches to stereotypes and stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 1-35). Hillsdale, NJ: Erlbaum.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Bulletin*, *84*, 191-215.
- Barclay, C. R., & DeCooke, P. A. (1988). Ordinary everyday memories: Some of the things of which selves are made. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered* (pp. 91-125). New York: Cambridge University Press.
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, *54*, 462-479.
- Bernieri, F. J., & Gillis, J. S. (2001). Judging rapport: Employing Brunswik's lens model to study interpersonal sensitivity. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 67-88). Mahwah, NJ: Lawrence Erlbaum.
- Biernat, M. (1995). The shifting standards model: Implications of stereotype accuracy for social judgment. In Y. T. Lee, L. Jussim, & C. R. McCauley (Eds.), *Stereotype accuracy* (pp. 87-114). Washington, DC: American Psychological Association.
- Block, J. (1993). Studying personality the long way. In D. C. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds.), *Studying lives through time: Personality and development* (pp. 9-41). Washington, DC: American Psychological Association.
- Block, J. (1965). *The challenge of response sets*. New York: Meredith Publishing.
- Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of Personality and Social Psychology*, *55*, 726-737.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley Interscience.
- Borgida, E., Rudman, L. A., & Manteufel, L. L. (1995). On the courtroom use and misuse of gender stereotyping research. *Journal of Social Issues*, *51*, 181-192.
- Brewer, M. B. (1988a). A dual process model of impression formation. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Advances in social cognition* (Vol. 1, pp. 1-36). Hillsdale, NJ: Erlbaum.
- Brewer, W. F. (1988b). Memory for randomly sampled autobiographical events. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered* (pp. 21-90). New York: Cambridge University Press.
- Brigham, J. C. (1971). Ethnic stereotypes. *Psychological Bulletin*, *76*, 15-38.
- Brod, S. E., & Ross, L. D. (1998). The role of stereotyping in overconfident social prediction. *Social Cognition*, *16*, 225-252.
- Brophy, J. (Ed.) (1998). *Advances in Research on Teaching* (Vol. 7). Greenwich, Connecticut: JAI Press.
- Brophy, J. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology*, *75*, 631-661.
- Brophy, J., & Good, T. (1974). *Teacher-student relationships: Causes and consequences*. New York: Holt, Rinehart, and Winston.
- Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago: University of Chicago Press.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand-McNally.
- Carmine, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Beverly Hills, CA: Sage.
- Cialdini, R. B. (1993). *Influence: Science and Practice* (3rd ed.). New York: HarperCollins.

- Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology*, *33*, 541-560.
- Claire, T., & Fiske, S. (1998). A systemic view of behavioral confirmation: Counterpoint to the individualist view. In C. Sedikides, J. Schopler, & C. A. Insko (Eds.), *Intergroup cognition and intergroup behavior* (pp. 205-231). Mahwah, NJ: Erlbaum.
- Clarke, R. B., & Campbell, D. T. (1955). A demonstration of bias in estimates of Negro ability. *Journal of Abnormal and Social Psychology*, *51*, 585-588.
- Cohen, C. E. (1981). Personal categories and social perception: Testing some boundaries of the processing effects of prior knowledge. *Journal of Personality and Social Psychology*, *40*, 441-452.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.
- Craik, K. H., & Appleyard, D. (1980). Streets of San Francisco: Brunswik's lens model applied to urban inference and assessment. *Journal of Social Issues*, *36*, 72-85.
- Cronbach, L. J. (1955). Processes affecting scores on understanding of others and assumed similarity. *Psychological Bulletin*, *52*, 177-193.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302.
- Darley, J. M., & Fazio, R. H. (1980). Expectancy-confirmation processes arising in the social interaction sequence. *American Psychologist*, *35*, 867-881.
- Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, *44*, 20-33.
- Dawes, & Robyn, M. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York: Free Press.
- Dawes, R. M. (1989). Statistical criteria for establishing a truly false consensus effect. *Journal of Experimental Social Psychology*, *25*, 1-17.
- Dawes, R. M. (1988). *Rational Choice in an Uncertain World*. San Diego, CA: Harcourt Brace Jovanovich.
- Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*, 571-582.
- Devine, P. (1995). *Prejudice and outgroup perception*. In A. Tesser (Ed.), *Advanced Social Psychology* (pp. 467-524). New York: McGraw-Hill.
- Diekmann, A. B., Eagly, A. H., & Kulesa, P. (2002). Accuracy and bias in stereotypes about the social and political attitudes of women and men. *Journal of Experimental Social Psychology*, *38*, 268-282.
- Dunning, D., & Sherman, D. A. (1997). Stereotypes and tacit inference. *Journal of Personality and Social Psychology*, *73*, 459-471.
- Dusek, J. (Ed.) (1985). *Teacher expectancies*. Hillsdale, NJ: Erlbaum.
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. San Diego, CA: Harcourt Brace Jovanovich.
- Eagly, A. H., & Diekmann, A. B. (1997). The accuracy of gender stereotypes: A dilemma for feminism. *International Review of Social Psychology*, *10*, 11-30.
- Eagly, A. H., Makhijani, M. G., Ashmore, R. D., & Longo, L. C. (1991). What is beautiful is good, but ...: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, *110*, 109-128.
- Eccles, J., & Wigfield, A. (1985). Teacher expectations and student motivation. In J. Dusek (Ed.), *Teacher expectancies* (pp. 185-226). Hillsdale, NJ: Erlbaum.

- Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and cues to deception. *Psychiatry*, 32, 88-105.
- Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology*, 29, 288-298.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality & Social Psychology*, 69, 1013-1027.
- Felson, R. B. (1984). The effect of self-appraisals of ability on academic performance. *Journal of Personality & Social Psychology*, 47, 944-952.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Festinger, L., Riecken, H. W., & Schachter, S. (1956). *When prophecy fails*. Minneapolis: University of Minnesota Press.
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 357-411). New York: McGraw-Hill.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, 23, 1-74.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.
- Fong, G. T., & Nisbett, R. E. (1991). Immediate and delayed transfer of training effects in statistical reasoning. *Journal of Experimental Psychology: General*, 120, 34-45.
- Funder, D. C. (2001). Personality. *Annual Review of Psychology*, 52, 197-221.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, 102, 652-670.
- Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception*. San Diego, CA: Academic Press.
- Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. *Psychological Bulletin*, 101, 75-90.
- Funder, D. C., & West, S. G. (1993). Consensus, self-other agreement, and accuracy in personality judgment: An introduction. *Journal of Personality*, 61, 457-476.
- Gage, N. L., & Cronbach, L. J. (1955). Conceptual and methodological problems in interpersonal perception. *Psychological Review*, 62, 411-422.
- Gergen, K. J. (1985). The social constructionist movement in modern psychology. *American Psychologist*, 40, 266-275.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin.
- Gilbert, D. T. (1995). Attribution and interpersonal perception. In A. Tesser (Ed.), *Advanced social psychology* (pp. 99-147). New York: McGraw-Hill.
- Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. New York: The Free Press.
- Goldman, W., & Lewis, P. (1977). Beautiful is good: Evidence that the physically attractive are more socially skillful. *Journal of Experimental Social Psychology*, 13, 125-130.
- Gottfredson, L. S. (1994). The science and politics of race-norming. *American Psychologist*, 49, 955-963.
- Gould, S. J. (1978). *The mismeasure of man*. New York: Norton.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464-1480.
- Hall, J. A., & Carter, J. D. (1999). Gender-stereotype accuracy as an individual difference. *Journal of Personality and Social Psychology*, 77, 350-359.

- Hamilton, D. L., Sherman, S. J., & Ruvolo, C. M. (1990). Stereotype-based expectancies: Effects on information processing and social behavior. *Journal of Social Issues*, 46, 35-60.
- Hastie, R., & Dawes, R. (2001). *Rational choice in an uncertain world*. Thousand Oaks, CA: Sage.
- Hastorf, A. H., & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology*, 47, 129-143.
- Hastie, R., & Rasinski, K. A. (1988). The concept of accuracy in social judgment. In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 193-208). Cambridge: Cambridge University Press.
- Hedges, S. M., Jandorf, L., & Stone, A. A. (1985). Meaning of daily mood assessments. *Journal of Personality and Social Psychology*, 48, 428-434.
- Herrnstein, R. J., & Murray, C. (1994). *The bell curve*. New York: The Free Press.
- Hoge, R. D., & Butcher, R. (1984). Analysis of teacher judgments of pupil achievement levels. *Journal of Educational Psychology*, 76, 777-781.
- Ickes, W. J. (1997). *Empathic accuracy*. New York: The Guilford Press.
- Ickes, W. (1993). Empathic accuracy. *Journal of Personality*, 61, 587-610.
- Jacoby, R., & Glauber, N. (1995). *The bell curve debate*. New York: Times Books.
- Jones, E. E. (1985). Major developments in social psychology during the past five decades. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (3rd ed., Vol. 1, pp. 47-107). New York: Random House.
- Jones, E. E. (1986). Interpreting interpersonal behavior: The effects of expectancies. *Science*, 234, 41-46.
- Jones, E. E. (1990). *Interpersonal perception*. New York: W. H. Freeman and Company.
- Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, 33, 1-27.
- Jost, J. T., & Kruglanski, A. W. (2002). The estrangement of social constructionism and experimental social psychology: History of the rift and prospects for reconciliation. *Personality and Social Psychology Review*, 6, 168-187.
- Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review*, 100, 109-128.
- Judd, C. M., Park, B., Ryan, C. S., Brauer, M., & Kraus, S. (1995). Stereotypes and ethnocentrism: Diverging interethnic perceptions of African American and White American youth. *Journal of Personality & Social Psychology*, 69, 460-481.
- Jussim, L. (1986). Self-fulfilling prophecies: A theoretical and integrative review. *Psychological Review*, 93, 429-445.
- Jussim, L. (1989). Teacher expectations: Self-fulfilling prophecies, perceptual biases, and accuracy. *Journal of Personality and Social Psychology*, 57, 469-480.
- Jussim, L. (1991). Social perception and social reality: A reflection-construction model. *Psychological Review*, 98, 54-73.
- Jussim, L. (2004). The goodness of judgment index. *Behavioral and Brain Sciences*, 27, 344-345.
- Jussim, L., & Harber, K. (In press). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*.
- Jussim, L., Harber, K., Crawford, J., Cain, T., & Cohen, F. (in press). *Social reality makes the social mind: Self-fulfilling prophecy, stereotypes, bias, and accuracy in social perception*. *Interaction Studies*.
- Jussim, L., Eccles, J., & Madon, S. J. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. *Advances in Experimental Social Psychology*, 29, 281-388.

- Jussim, L., McCauley, C. R., & Lee, Y. T. (1995). Why study stereotype accuracy and inaccuracy? In Y. T. Lee, L. Jussim, & C. R. McCauley (Eds.), *Stereotype accuracy: Toward appreciating group differences* (pp. 3-27). Washington, DC: American Psychological Association.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.
- Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation* (Vol. 15). Lincoln: University of Nebraska Press.
- Kelly, G. A. (1955). *The psychology of personal constructs* (Vol. 1). New York: W. W. Norton.
- Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York: Guilford.
- Kenny, D. A., & Acitelli, L. K. (2001). Accuracy and bias in the perception of the partner in a close relationship. *Journal of Personality and Social Psychology*, 80, 439-448.
- Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin*, 102, 390-402.
- Kenny, D. A., Kashy, D., & Bolger, N. (1998). Data analysis in social psychology. In D. Gilbert, S. I. Fiske, & J. Lindzey (Eds.), *Handbook of Social Psychology* (4th ed.). New York: McGraw Hill.
- Koriat, A., Goldsmith, M., & Panksy, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology*, 51, 481-537.
- Krueger, J. (1996). Probabilistic national stereotypes. *European Journal of Social Psychology*, 26, 961-980.
- Krueger, J., & Funder, D. C. (2004). Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences*, 27, 313-367.
- Krueger, J. I., Hasman, J. F., & Acevedo, M. (2003). Perceptions of trait typicality in gender stereotypes: Examining the role of attribution and categorization processes. *Personality and Social Psychology Bulletin*, 29, 108-116.
- Kruglanski, A. (1989). The psychology of being "right": The problem of accuracy in social perception and cognition. *Psychological Bulletin*, 106, 395-409.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480-498.
- Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, 103, 284-308.
- LaPiere, R. T. (1934). Attitudes versus actions. *Social Forces*, 13, 230-237.
- LaPiere, R. T. (1936). Type-rationalizations of group antipathy. *Social Forces*, 15, 232-237.
- Latane, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help*. New York: Appleton-Century-Crofts.
- Lee, Y. T., Jussim, L., & McCauley, C. R. (Eds.) (1995). *Stereotype Accuracy: Toward Appreciating Group Differences*. Washington, DC: American Psychological Association.
- Locksley, A., Borgida, E., Brekke, N., & Hepburn, C. (1980). Sex stereotypes and social judgment. *Journal of Personality and Social Psychology*, 39, 821-831.
- Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology*, 47, 1231-1243.
- Lord, C. 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.
- Madon, S., Guyl, M., Spoth, R. L., Cross, S. E., & Hilbert, S. J. (2003). The self-fulfilling influence of mother expectations on children's underage drinking. *Journal of Personality and Social Psychology*, 84, 1188-1205.

- Madon, S. J., Jussim, L., Keiper, S., Eccles, J., Smith, A., & Palumbo, P. (1998). The accuracy and power of sex, social class and ethnic stereotypes: Naturalistic studies in person perception. *Personality and Social Psychology Bulletin*, 24, 1304-1318.
- Malkiel, B. G. (1999). *A random walk down Wall Street*. New York: Norton.
- McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, 90, 215-238.
- McCauley, C. (1995). Are stereotypes exaggerated? A sampling of racial, gender, occupational, academic, and political stereotypes. In Y. T. Lee, L. Jussim, & C. R. McCauley (Eds.), *Stereotype accuracy: Toward appreciating group differences* (pp. 215-243). Washington, DC: American Psychological Association.
- McCauley, C., Jussim, L., & Lee, Y. T. (1995). Stereotype accuracy: Toward appreciating group differences. In Y. T. Lee, L. Jussim, & C. R. McCauley (Eds.), *Stereotype accuracy: Toward appreciating group differences* (pp. 293-312). Washington, DC: American Psychological Association.
- McCauley, C., & Stitt, C. L. (1978). An individual and quantitative measure of stereotypes. *Journal of Personality and Social Psychology*, 36, 929-940.
- McCauley, C., Stitt, C. L., & Segal, M. (1980). Stereotyping: From prejudice to prediction. *Psychological Bulletin*, 87, 195-208.
- McCauley, C., & Thangavelu, K. (1991). Individual differences in strength of sex-occupation stereotyping: Reliability, validity, and some correlates. *Social Psychology Quarterly*, 54, 195-208.
- McCauley, C., Wright, M., & Harris, M. E. (2000). Diversity workshops on campus: A survey of current practice at U.S. colleges and universities. *College Student Journal*, 34, 100-114.
- McNatt, D. B. (2000). Ancient Pygmalion joins contemporary management: A meta-analysis of the result. *Journal of Applied Psychology*, 85, 314-322.
- Meehl, P. E. (1990). Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. *Psychological Inquiry*, 1, 108-141.
- Merton, R. K. (1948). The self-fulfilling prophecy. *Antioch Review*, 8, 193-210.
- Milgram, S. (1974). *Obedience to authority*. New York: Harper & Row.
- Miller, D. T., & Turnbull, W. (1986). Expectancies and interpersonal processes. *Annual Review of Psychology*, 37, 233-256.
- Murray, S. L., Holmes, J. G., & Griffin, D. W. (1996). The self-fulfilling nature of positive illusions in romantic relationships: Love is not blind, but prescient. *Journal of Personality and Social Psychology*, 71, 1155-1180.
- Myers, D. G. (1999). *Social psychology* (6th ed.). New York: McGraw-Hill.
- Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77-101.
- Nelson, T. (2002). *The psychology of prejudice*. Boston: Allyn & Bacon.
- 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. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). *Stereotyping and social reality*. Cambridge, MA: Blackwell.
- Olson, J. M., Roesse, N. J., & Zanna, M. P. (1996). Expectancies. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 211-239). New York: Guilford Press.

- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes*. San Diego, CA: Academic Press.
- Paulhus, D. L. (2002). Socially desirable responding: Evolution of a construct. In H. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement*. Mahwah, NJ: Erlbaum.
- Paunonen, S. V. (1991). On the accuracy of ratings of personality by strangers. *Journal of Personality and Social Psychology*, *61*, 471-477.
- Pickering, M. (2001). *Stereotyping*. New York: Palgrave.
- Pinderhughes, E. (1989). *Understanding race, ethnicity, and power: The key to efficacy in clinical practice*. New York: The Free Press.
- Räikkönen, K., Matthews, K. A., Flory, J. S., Owens, J. F., & Gump, B. B. (1999). Effects of optimism, pessimism, and trait anxiety on ambulatory blood pressure and mood during everyday life. *Journal of Personality and Social Psychology*, *76*, 104-113.
- Raudenbush, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy inductions: A synthesis of findings from 18 experiments. *Journal of Educational Psychology*, *76*, 85-97.
- Redding, R. E. (2001). Sociopolitical diversity in psychology: The case for pluralism. *American Psychologist*, *56*, 205-215.
- Reynolds, D. J., Jr., & Gifford, R. (2001). The sounds and sights of intelligence: A Lens Model channel analysis. *Personality and Social Psychology Bulletin*, *65*, 187-2000.
- Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (Eds.) (1991). *Measures of personality and social psychological attitudes*. San Diego, CA: Academic Press.
- Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (Eds.) (1999). *Measures of political psychological attitudes*. San Diego, CA: Academic Press.
- Rosenhan, D. L. (1973). On being sane in insane places. *Science*, *179*, 250-258.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectations and student intellectual development*. New York: Holt, Rinehart, and Winston.
- Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *The Behavioral and Brain Sciences*, *3*, 377-386.
- Ross, L., & Nisbett, R. E. (1991). *The person and the situation*. New York: McGraw-Hill.
- Ryan, C. S. (2002). Stereotype accuracy. *European Review of Social Psychology*, *13*, 75-109.
- Ryan, C. S. (1996). Accuracy of black and white college students' in-group and out-group stereotypes. *Personality and Social Psychology Bulletin*, *22*, 1114-1127.
- Ryan, C. S. (1995). Motivations and the perceiver's group membership: Consequences for stereotype accuracy. In Y. T. Lee, L. Jussim, & C. R. McCauley (Eds.), *Stereotype accuracy* (pp. 189-214). Washington, DC: American Psychological Association.
- Ryan, C. S., & Bogart, L. M. (2001). Longitudinal changes in the accuracy of new group members' in-group and out-group stereotypes. *Journal of Experimental Social Psychology*, *37*, 118-133.
- Ryan, C., Park, B., & Judd, C. M. (1996). Assessing stereotype accuracy: Implications for understanding the stereotyping process. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping*. New York: Guilford.
- Schmitt, N., Noe, R. A., & Gottschalk, R. (1986). Using the lens model to magnify raters' consistency, matching, and shared bias. *Academy of Management Journal*, *29*, 130-139.
- Schneider, D. J., Hastorf, A. H., & Ellsworth, P. C. (1979). *Person perception* (2nd ed.). Reading, MA: Addison-Wesley.
- Seligman, & Martin, E. P. (1995). The effectiveness of psychotherapy: The Consumer Reports study. *American Psychologist*, *50*, 965-974.

- Sidanius, J., & Pratto, F. (1999). *Social dominance: an intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.
- Snyder, M. (1984). When belief creates reality. *Advances in Experimental Social Psychology*, *18*, 247-305.
- Snyder, M. (1992). Motivational foundations of behavioral confirmation. *Advances in Experimental Social Psychology*, *18*, 247-305.
- Snyder, M., & Stukas, A. A., Jr. (1998). Interpersonal processes: The interplay of cognitive, motivational, and behavioral activities in social interaction. *Annual Review of Psychology*, *50*, 273-303.
- Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, *35*, 656-666.
- Spitz, H. H. (1999). Beleaguered Pygmalion: A history of the controversy over claims that teacher expectancy raises intelligence. *Intelligence*, *27*, 199-234.
- Stangor, C. (1995). Content and application inaccuracy in social stereotyping. In Y. T. Lee, L. Jussim, & C. R. McCauley (Eds.), *Stereotype accuracy* (pp. 275-292). Washington, DC: American Psychological Association.
- Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*, *111*, 42-61.
- Stinson, L., & Ickes, W. (1992). Empathic accuracy in the interactions of male friends versus male strangers. *Journal of Personality and Social Psychology*, *62*, 787-797.
- Swann, W. B. (1984). Quest for accuracy in person perception: A matter of pragmatics. *Psychological Review*, *91*, 457-477.
- Swim, J. K. (1994). Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes. *Journal of Personality and Social Psychology*, *66*, 21-36.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193-210.
- Trouilloud, D., Sarrazin, P., Martinek, T., & Guillet, E. (2002). The influence of teacher expectations on students' achievement in physical education classes: Pygmalion revisited. *European Journal of Social Psychology*, *32*, 1-17.
- Tversky, A., & Kahneman, D. (1983). Extensions versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293-315.
- U. S. Census (2002). *Money income in the United States: 2001* (Web address: <http://www.census.gov/prod/2002pubs/p60-218.pdf>).
- West, C., & Anderson, T. (1976). The question of preponderant causation in teacher expectancy research. *Review of Educational Research*, *46*, 613-630.
- Williams, T. (1976). Teacher prophecies and the inheritance of inequality. *Sociology of Education*, *49*, 223-236.
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101-126.
- Wolsko, C., Park, B., Judd, C. M., & Wittenbrink, B. (2000). Framing interethnic ideology: Effects of multicultural and color-blind perspectives on judgments of groups and individuals. *Journal of Personality and Social Psychology*, *78*, 635-654.
- Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology*, *10*, 109-120.
- Wright, J. C., & Drinkwater, M. (1997). Rationality vs. accuracy of social judgment. *Social Cognition*, *15*, 245-273.