On the Accuracy of Personality Judgment: A Realistic Approach

David C. Funder
University of California, Riverside

The "accuracy paradigm" for the study of personality judgment provides an important, new complement to the "error paradigm" that dominated this area of research for almost 2 decades. The present article introduces a specific approach within the accuracy paradigm called the Realistic Accuracy Model (RAM). RAM begins with the assumption that personality traits are real attributes of individuals. This assumption entails the use of a broad array of criteria for the evaluation of personality judgment and leads to a model that describes accuracy as a function of the availability, detection, and utilization of relevant behavioral cues. RAM provides a common explanation for basic moderators of accuracy, sheds light on how these moderators interact, and outlines a research agenda that includes the reintegration of the study of error with the study of accuracy.

Believe truth! Shun error!—these, we see, are two materially different laws; and by choosing between them we may end by coloring differently our whole intellectual life. (James, 1897/1915)

Judgments of personality are attempts to identify the psychological properties of people, such as personality traits, that help to explain what they have done in the past and to predict what they will do in the future (Funder, 1991). For example, one person might judge the degree to which another is critical, dependable, or energetic. The judgment might be used to explain why he or she insulted an acquaintance or to predict whether he or she will work at a job reliably and well. Judgments like these are not rare in the lab or in daily life and frequently have been the focus of psychological research.

For almost two decades, the dominant, implicit motto of this research surely has been "Shun error!" The detection of error has constituted an important approach toward the study of how people judge each others' personalities, as well as of human judgment more generally (e.g., Kahneman & Tversky, 1973; Ross & Nisbett, 1991). A prolific, ingenious, and influential program of research has exhaustively detailed characteristic shortcomings of human judgment (cf. Ross, 1977, and Nisbett & Ross, 1980), generated suggestions about how to mitigate some of these shortcomings, and yielded insights into processes or "heuristics" of human judgment that can produce error (Funder, 1987; Jones, 1985).

Still, successful as it has been, the "error paradigm" can tell no more than half of the story. The avoidance of error is not

quite the same thing as the achievement of accuracy, and explanations of how errors arise shed relatively little light on how correct judgments are ever made. As James (1897/1915) pointed out nearly a century ago, the shunning of error needs to be complemented by a more positively oriented seeking after truth.

Just this sort of complement has begun to emerge from an alternative approach to the study of personality judgment that could be called the "accuracy paradigm" (e.g., Funder, 1987; Kenny, 1994; Kruglanski, 1989; Swann, 1984; for a collection of representative research, see Funder & West, 1993b). Followed by an increasing number of researchers in recent years, this new paradigm focuses not so much on what judges cannot do but on what they can do and the circumstances under which they can do it. Although the accuracy paradigm seems—and is—relatively optimistic compared with the error paradigm, it is not really Panglossian. Judgmental mistakes obviously are frequent in daily life. But the accuracy paradigm incorporates a recognition that progress in this area requires examination of when and how people are correct as well as when and how they are mistaken.

Varieties of Accuracy

The modern accuracy paradigm comprises three major variants: the pragmatic, constructivist, and realistic approaches (Funder & West, 1993a). The pragmatic approach views personality judgments as necessary tools for social living and evaluates their accuracy in terms of their practical value (McArthur & Baron, 1983; Swann, 1984). A judgment is viewed as accurate if it leads to successful social interaction. The constructivist approach views personality judgments as social constructions and evaluates their accuracy in terms of consensus, or agreement between judges (Kruglanski, 1989). In its extreme form, constructivism implies that if judges agree, they must be considered accurate, because no other criterion is available even in principle. In its less extreme form, research focuses exclusively or nearly exclusively on the factors that promote and in-

hibit interjudge agreement independent of accuracy (Kenny, 1994).

Both of these approaches provide useful insights, and the research that emanates from them constitutes an important part of the accuracy paradigm. However, the present article is organized around and will argue specifically for a third approach, to be called the Realistic Accuracy Model (RAM). RAM begins with a general and consequential assumption that sets it apart from the error paradigm as well as the other two approaches to accuracy. Its basic, "neo-Allportian" assumption is that personality traits are real characteristics of individuals (Allport, 1937; Funder, 1991, 1993b).

This assumption yields two implications. The first is that the accuracy of personality judgment is an extremely complex matter. It goes beyond relatively convenient operational definitions and into complex issues concerning the construct validity of personality traits (Cronbach & Meehl, 1955; Ozer, 1989). The evaluation of accuracy requires consideration of the widest range of data that an investigator can gather and moves the study of social judgment off the familiar territory of social psychology onto some of the traditional turf of personality psychology.

The second implication is that accuracy in personality judgment is a joint product of the attributes and behavior of the target as well as of the observations and perspicacity of the judge. In other words, accuracy stems from the relevance, availability, detection, and utilization of behavioral cues. Other approaches to error and accuracy do not yield this implication because they pay much more attention to the judge than to the person who is judged. The assumption that traits are real determines the distinctive way research within RAM seeks data, evaluates accuracy, and organizes results.

The Importance of Accuracy

Not all judgments of personality concern traits, but many certainly do. Consider the typical letter of recommendation. Trait constructs abound (e.g., "the candidate is conscientious, energetic, insightful, and pleasant"), and if these constructs have any meaning at all, then their application might be right, wrong, or someplace in between.[1] Or, a college student home on vacation might be asked by her mother to describe her new roommate. "She is friendly, but sort of sloppy," says the student, "and very hardworking." Presumably, these terms are meant to describe something real. Or, to be more precise, each of these terms intends to describe two real things: a pattern of behavior and an inferred attribute of the person who performed it (Funder, 1991).

Sometimes personality judgments are made and communicated in an attempt to predict the future behavior of the person whose attributes are described. For example, the reader of the letter of recommendation may have to make a decision about whether to admit or hire the person in question. The basis of this decision is in large part a prediction of the person's behavior in school or on the job that the reader bases on the traits described in the letter.

But very often the goal of trait description is as much intrinsic as it is pragmatic. People are naturally curious about each other,

and the college student's mother may want to know about who her daughter is living with, not really to predict the roommate's behavior, but just because she cares. The nature of much interpersonal gossip demonstrates how interested people are in judging and communicating about each others' salient characteristics, sometimes for what seems like no particular reason, which is precisely what shows the interest to be intrinsic.

The accuracy of personality judgment is also important for methodological, theoretical, and philosophical reasons (Funder & West, 1993a). From a methodological perspective, human judgments of individual characteristics are an important source of data for personality, developmental, and clinical psychology. Many studies ask informants, who may be lay peers or clinical assessors, to use a set of rating scales or a Q-sort deck to distill and describe their impressions (e.g., Bem & Funder, 1978; Block, Block, & Harrington, 1974; Block & Robins, 1993; Colvin, Block, & Funder, 1995; Funder & Colvin, 1991). These judgments are then used as data about the personality attributes of the subjects they describe. The quality of these data and the validity of the conclusions drawn from them depend critically on the accuracy of the judgments (Funder, 1993a).

From a theoretical perspective, the accuracy of personality judgment is intertwined with the issue of how personality is manifest in behavior. The realistic assumption of RAM implies that to understand when and how a trait can be inferred, one has to understand when and how that trait influences what people do. This latter issue is a traditional concern of personality psychology, and for that reason accuracy research is as relevant to that field as it is to social psychology, the more usual home for research on interpersonal judgment in recent years (Allport, 1937; Funder, 1993b).

From a philosophical perspective, the key questions about perception and judgment pertain to the knowledge people possess, the way this knowledge is obtained, and the degree to which it reflects the true state of nature (Hastie & Rasinski, 1988; Kruglanski & Ajzen, 1983). RAM's realistic approach specifically addresses the connection between social perception and the psychological reality that it presumes to lie beneath.

From Accuracy to Error to Accuracy Again

The First Wave of Accuracy Research

The history of research on personality judgment over the past 60 years shows an interest in accuracy that has waxed, waned, and then recently waxed again.[2] Certainly, from a lay, naive, or commonsensical point of view, the most obvious question that might be asked about a personality judgment is, "Is it right or wrong?" It was only natural, therefore, for research on personality judgment to begin, in the mid-1930s, with a focus on ac-

---

[1] A radical constructivist view would regard words like these as essentially meaningless in that they are not seen as referring to anything other than arbitrary ideas in the mind of the perceiver. Such an approach stops the analysis of accuracy before it even starts and therefore must be rejected here.

[2] For much more detailed renditions of this history, see Funder (1987), Funder and West (1993a), and Kenny (1994).

curacy (e.g., Allport, 1937, chapter 18). A particular concern was the search for the "good judge" of personality, the individual difference characteristics that would make a person more likely to be an accurate judge of others (Estes, 1938; Taft, 1955).

A lively research tradition organized around this topic came to an abrupt halt in the mid-1950s, however, for at least three reasons that can now be identified. The first is that the search for the good judge proved by any standard disappointing. Personality correlates of judgmental ability were weak and changeable across studies, and judgmental ability itself often was inconsistent. The same person who was a good judge in one context might be a poor judge in another context that was only slightly different (Cline & Richards, 1960; Crow & Hammond, 1957; Schneider, Hastorf, & Ellsworth, 1979).

A second cause for the falloff in interest in accuracy was the publication, in 1955, of a set of severe critiques of the methods used by almost all the research on the topic reported up to that time (Cronbach, 1955; Gage & Cronbach, 1955; see also Hastorf & Bender, 1952). Writing a pair of difficult articles using unconventional statistical notation, Lee Cronbach nonetheless succeeded at convincing his colleagues that nearly all the accuracy research he reviewed was nearly meaningless. The basic reason was that the numbers used to reflect interjudge agreement—the usual criterion for accuracy then as now—were potentially influenced by several more-or-less artifactual factors including response sets, actual and assumed similarity between judge and target, and so forth.

Cronbach's intention may have been to improve accuracy research rather than to shut it down, but the latter is very nearly what happened. Both the complex statistical adjustments recommended in a precomputer era and the large amount of data gathering that Cronbach argued was required apparently proved daunting to many investigators. As Schneider et al. (1979) noted, the topic suddenly "lost some of its intuitive charm" (p. 222). For more than two decades (until approximately 1980),[3] very few studies of accuracy followed Cronbach's critique.

The Focus on Cognitive Process

A third reason for the falloff in accuracy research in the mid-1950s was the growth of an alternative paradigm within the study of what became called *person perception* (e.g., Bruner & Tagiuri, 1954; Tagiuri & Petrullo, 1958). This paradigm moved the study of interpersonal judgment into the laboratory, where subjects could be induced to make judgments of artificial stimuli such as lists of trait words rather than of real people. Asch (1946) and others discovered that such studies could reveal interesting aspects of the way information is combined on the way to personality judgments, without requiring any concern with (and without providing any information about) the social content or accuracy of these judgments.

Consider one of the pioneering studies in this tradition, the one by Asch (1946). In cleverly designed experiments, Asch presented subjects with bits of descriptions of hypothetical stimulus persons (e.g., a list of trait words). He showed how the degree to which a judgment was affected by a given piece of

information depended on whether the information was presented first, in the middle, or last and whether it seemed consistent or contradictory relative to the rest of the information presented. This was and remains important work, but notice how its conclusions are not specifically relevant to personality. Analogous findings could be—and are—found concerning other kinds of information leading to other kinds of judgment (cf. J. R. Anderson, 1990; Glanzer & Cunitz, 1966; Koppenaal & Glanzer, 1990). The research is in that sense "content-free": It addresses the cognitive process rather than social substance of judgment, and this process occurs wholly in the mind of the judge rather than in the interpersonal world.

By the same token, such research and that which it inspired say little or nothing about the variables that might affect accuracy. Asch's (1946) "primacy effect," for example, will enhance accuracy if the content of the information presented first happens to be diagnostic of what is judged, and the very same effect will lessen accuracy if the information is undiagnostic or misleading.

However, accuracy was never the point of this kind of research on person perception. Asch, for one, never claimed that his studies had accuracy implications. A later but equally prominent practitioner of cognitively oriented research on person perception, E. E. Jones, explicitly commented about how studies in the Aschian style, including his own, "solved the accuracy problem by bypassing it" (Jones, 1985, p. 77). In context, Jones was not acknowledging any deficiency. Rather, he viewed it as salutary that research in person perception had found a way around the difficult and perhaps intractable issues involved in the evaluation of judgmental accuracy. Over the next decades, a great deal of research on "attribution theory" and, later, "social cognition" grew out of this approach. Nearly all of it used experimentally manipulated, artificial social stimuli that were ideally suited for testing models of some of the fine cognitive processes involved in person perception. None of it said very much about accuracy one way or the other (e.g., Fiske & Taylor, 1991; Schneider et al., 1979).

The Rise of Error

What was banished from the front door began creeping in slowly through the back door, however, with the gradual rise of research on judgmental error. This approach originated within

---

[3] The literature on industrial/organizational (I/O) psychology provides an important exception, notably as it has appeared in the pages of the *Journal of Applied Psychology*. Perhaps because of its relative isolation from mainstream personality and social psychology research, or perhaps because of its pressing concerns with applied issues of personnel selection and placement, I/O psychology maintained a steady interest in accuracy throughout the 1950s and 1960s and into the present day (e.g., Borman, 1977; Gifford, Fan Ng, & Wilkinson, 1985; Hollander, 1957, 1965; Kane & Lawler, 1978; Lewin & Zwany, 1976; Mayfield, 1972; Paunonen & Jackson, 1987; Schmidt, Hunter, Croll, & McKenzie, 1983; Waters & Waters, 1970). The research by Jackson and his coworkers maintained a consistent interest and active research program in accuracy that also addressed clinical and developmental issues (e.g., MacLellan & Jackson, 1985; Strasburger & Jackson, 1977).

cognitive psychology (e.g., Kahneman & Tversky, 1973) and quickly was extended to the study of person perception (e.g., Nisbett & Ross, 1980). Within both cognitive and social psychology, the strategy was to posit hypothetical models of the process of judgment, which were tested by examining how subjects responded to artificial or contrived stimuli. Some of these models—such as Bayesian statistical inference, expected utility theory, and the analysis of variance-based "Kelley cube"—began to be ascribed normative status. Rather than just hypotheses about how people might think, these models were treated as prescriptions for how people should think (Gigerenzer, 1991a). Judgments that differed from these prescriptions were seen as errors, and the ease and frequency with which such errors could be demonstrated were interpreted as indicating not only that the models imperfectly described human judgment but that human judgment itself was imperfect.

For a considerable period the error paradigm held the field nearly alone in the study of personality judgment. Research identified a large number of errors, described their underlying processes or heuristics, and was widely cited (Ross & Nisbett, 1991). As this work became increasingly popular, its tight focus on error produced a widespread impression that, in general, the accuracy of human judgment is poor (Funder, 1987; Lopes, 1991). The unquestioned dominance of this approach was not to last, however. By the late 1980s, the error paradigm had received its own share of criticism.

Critiques of Error

Modern critiques of the error paradigm emphasized three points. First, at least a few—and perhaps more than a few—of the errors identified by this research seem on close examination somewhat artificial. They are displayed in experimental contexts carefully designed to evoke them, whereas their frequency and meaning in realistic contexts often goes unexamined (Gigerenzer, 1991b). Some—such as the "fundamental attribution error," the tendency to see behavior as caused by personal dispositions even in the absence of evidence to that effect—may, as often as not, produce correct judgments in realistic contexts (Funder, 1987, 1993a; Jussim, 1993).

This observation leads to a second point: Little if any research has shown that the elimination of errors actually improves judgmental accuracy. Such evidence that does exist indicates rather the reverse (Block & Funder, 1986; Borman, 1975, 1979). Eliminating the "halo effect," for example, has been shown to make judgments of real individuals less accurate (Bernardin & Pence, 1980). This is probably because socially desirable traits usually do co-occur, making the inference of one such trait from the observation of another—the halo effect— a practice that ordinarily enhances accuracy (Funder & West, 1993a).

A third critical point is that, whatever its intrinsic scientific merits, research emanating from the error paradigm sometimes has been guilty of potentially consequential rhetorical excess. Writers have exaggerated the human propensity toward mistakes and underestimated accomplishments of the human judge (see Funder, 1993a; Gigerenzer, 1991b, 1991c; Lopes, 1991). The result is a view of human nature that is probably unduly pessimistic.

But William James's discomfort with a one-sided strategy of "shunning error," expressed nearly a century before, had a deeper basis than the problems identified by these modern critiques. James concluded that after one had successfully identified and eliminated every possible error in one's judgment, one would be left with—nothing! Although James was well aware of the necessity for critical examination of evidence and the elimination of error, he argued that it was necessary also to put some energy into seeking out ideas that one can believe and to have a certain sympathy for the means by which belief can be achieved. The realization that this latter approach is fundamentally different from the detection of error sets the stage for the development of the accuracy paradigm.

The New Accuracy Paradigm

The Roots of the New Paradigm

An alternative approach to the appraisal of accuracy in personality judgment began to develop in the early 1980s. The new paradigm arose in part as a reaction to and correction of the error paradigm and some of its perceived shortcomings. It also can in an important sense be considered a rebirth of the older research on accuracy from the 1940s and 1950s. But, this time around, it has incorporated a greater methodological and conceptual sophistication aimed at helping it to avoid some of the mistakes of the past (Funder & West, 1993a; Kenny, 1993, 1994).

The accuracy paradigm also has deeper intellectual roots in the work of Egon Brunswik (1956) and J. J. Gibson (1979). Both of these earlier investigators had in the field of physical perception (and in Brunswik's case, the study of judgment) long argued for a focus on the connections between perception and reality in real-life settings. Their descendants in personality and social psychology maintain that personality judgment should whenever possible be studied using targets who are real people observed in real settings and that criteria for accuracy should be drawn from the social environment. These criteria include consensus, self-other agreement, and behavioral prediction (Funder & West, 1993a; Kenny, 1994).

The Development of Research on Accuracy

The very existence of accuracy. Modern research on accuracy can be seen as developing through three overlapping phases. In the first phase, which may be very nearly over, accuracy research put most of its effort into restoring the plausibility of the idea that human judgments of personality could manifest any reasonable degree of validity at all. This seemed necessary because of the impression of almost universal inaccuracy that some readers derived from research on error. In this phase, investigators reported studies showing how personality traits affect behavior and how laypersons can make judgments of such traits that manifest both interjudge agreement and predictive validity (e.g., Cheek, 1982; Funder, 1980a, 1980b, 1982; Kenrick & Stringfield, 1980; Moskowitz & Schwarz, 1982). These two kinds of criteria—agreement and behavioral prediction—

became the basis of accuracy research that followed over the next decade (Funder, 1987; Funder & West, 1993a).

Moderators of accuracy. The second stage of accuracy research, still in progress, is to discover the moderator variables that make accurate judgments of personality more and less likely. Moving beyond "Are personality judgments ever accurate?" accuracy research has begun to address the next question, "When are personality judgments accurate?"

Research already has examined numerous variables that might affect accuracy. These variables have been quite diverse, but in previous writings I have proposed that they can be organized into four broad categories: (a) good judge, the possibility that some individuals might be better judges of personality than others; (b) good target, the possibility that some individuals might be more easily judged than others; (c) good trait (or behavior), the possibility that some traits (and therefore some behaviors) might be easier to judge (or to predict) accurately than others; and (d) good information, the possibility that more or certain kinds of information might make accurate judgments more likely. During the past few years, this four-part framework has served to organize numerous data analyses in our research program as well as a convenient way of summarizing principal empirical findings from our lab and elsewhere (see Funder, 1993a, for a review; see Allport, 1937, for a related framework). Recent research has steadily accumulated evidence as to the importance of accuracy for all but the first of these moderator variables. Some of this evidence will be reviewed later in this article.

The process of accurate judgment. The four-variable approach to moderators of accuracy has proved itself a useful way to organize a literature that otherwise might have seemed rather chaotic. In addition, it continues to generate research. However, in the broader scheme of things, this list of moderators constitutes just a necessary beginning. Having established that judgments of personality are sometimes accurate, and having begun to document when they are accurate, the necessary third stage for accuracy research is to begin to explain how they become accurate. Such an explanation is the purpose of RAM.

The Realistic Accuracy Model

The first theoretical claim of RAM, that personality traits are real, has two implications. First, it implies a particular approach to the evaluation of accuracy that requires the use of an especially broad range of criteria. Second, it implies a description of the path between personality and its accurate judgment that leads through the availability, detection, and utilization of relevant behavioral cues. This description allows diverse moderators of accuracy already identified to be explained in terms of a relatively few common factors, suggests new moderators that might be found, and outlines how moderators of accuracy interact.

Criteria for Accuracy

A distinctive attribute of the accuracy paradigm is that it uses external evidence to indicate the degree to which a judgment might correctly reflect reality. For the study of personality judgment, this strategy becomes possible only when the stimuli to

judgment are real people in realistic settings, that is, when there is a reality to be correctly reflected. All three of the basic approaches to accuracy follow this research strategy.

Constructivist and pragmatic approaches. The pragmatic point of view, described by Swann (1984), is that an accurate judgment is one that allows the judge to function well in his or her social environment (see also McArthur & Baron, 1983). Application of this criterion requires the investigator to gather data about the judgments that subjects make in their natural social environments and to assess how well these individuals are faring.

The constructivist point of view, discussed by Kruglanski (1989), is that because reality cannot be known apart from human perceptions of it, accuracy must be considered a function of the collective point of view of a community of judges. Application of this criterion requires that subjects' judgments of actual acquaintances be compared with each other and their consensus assessed.

Both of these approaches step outside of and complement the error paradigm by attending to theoretical considerations and sources of data that go beyond the cognitive process by which judgments are made. Still, compared with a realistic approach, each has the net effect of narrowing what accuracy means and simplifying how it can be evaluated. For the pragmatic approach, the judge's social success provides a sufficient criterion, and for the constructivist approach, consensus among judges is definitive. For neither approach is truth in any broader or more abstract sense necessarily relevant; accuracy is not viewed as dependent on any properties that the target of judgment actually has. Nearly all the focus is on the judge, not the judged.

The realistic approach followed by RAM, in contrast, does not consider either pragmatic success or consensus to provide sufficient criteria for accuracy. Realism requires a broader range of information that, taken together, might indicate what the target of judgment is actually like. In this way, RAM attends to the object of judgment as well as to the judge and as a result makes even greater demands on a prospective investigator.

RAM's realistic approach. RAM evaluates accuracy according to a strategy consistent with a postpositivist philosophy of science variously called *critical realism* (Cook & Campbell, 1979) and *pancritical rationalism* (Rorer, 1991). RAM's approach also resembles the "probabilistic functionalism" of Egon Brunswik (1956). Simply put, the philosophical position underlying RAM is that although truth indeed exists, there is no sure pathway to it. There is only a wide variety of alternative pathways, each of which is extremely unsure.

A great temptation under these circumstances is not to venture on the journey at all. Because truth can never be known with certainty, why concern oneself with truth? There are plenty of more tractable issues that can generate useful research.[4] To echo the words of Jones (1985), one can indeed "solve the accuracy issue by bypassing it" (p. 77).

This "search for the key where the light is best" is not the Jamesian solution, nor is it the approach of RAM. Instead, RAM pre-

---

[4] This is friendly advice that I have received many times with respect to research on accuracy.

scribes that a portrayal of the actual psychological attributes of the target person be sought in combination with a wide range of information and the accuracy of a judgment of the target's personality be evaluated in terms of its congruence with this portrayal. The information about the target can and should take many forms. It might include self-ratings of personality; inventory scores; ratings by knowledgeable informants; direct observations of the target person's behavior (e.g., in a laboratory); and reports of the person's behavior in daily life, perhaps gathered using diary or "beeper" methods (Csikszentmihalyi & Larson, 1992; Funder, 1993a; Spain, 1994).

Multifaceted criteria. This multifaceted approach to the evaluation of accuracy is the fundamental attribute of RAM's empirical program. Its purpose is to transcend some of the shortcomings inherent in simple operational definitions or other single criteria.

For example, consider interjudge agreement. As Kenny (1994) noted, if a group of judges were to be perfectly accurate, they would necessarily manifest perfect agreement. Short of that, ample evidence indicates that the same variables one would expect to increase accuracy actually do increase interjudge agreement. It seems reasonable to assume that knowing one's target better, judging visible as opposed to hidden traits, and judging neutral as opposed to evaluatively loaded traits would all improve accuracy. Each of these circumstances does increase agreement (Funder & Colvin, 1988, in press; Funder & Dobroth, 1987; Funder, Kolar, & Blackman, 1995; John & Robins, 1994a). So the use of interjudge agreement as one criterion for accuracy is far from unreasonable.

However, this criterion is not sufficient. Many people can agree and still all be wrong. Ethnic prejudice and the practice of genocide provide obvious examples. Whereas a constructivist approach claims that there exist no criteria for accuracy beyond consensus, RAM's realistic approach insists that diverse evidence besides and independent of interjudge agreement also be considered relevant.

The other principal source of evidence for the accuracy of personality judgment is behavioral prediction (Funder, 1987). The idea is if a judgment of personality can be used to predict the behavior of the person who is judged, then the judgment must be in some sense accurate. This criterion is also less than perfect. For example, one can occasionally make a correct prediction for the wrong reasons, and sometimes one might make a correct prediction without inferring a trait at all, based on unmediated extrapolation from past experience.

The multicriterial approach of RAM provides a measure of protection against the first potential problem. Whereas a pragmatic approach to accuracy might consider a single correct behavioral prediction or perhaps a few useful predictions to be sufficient to establish accuracy, from RAM's perspective other information must be considered as well. If an incorrect trait judgment should somehow happen to lead to a correct prediction, this other information would provide the tip-off. In such a case, behavioral predictions in different contexts and the opinions of knowledgeable informants and of the target person himself or herself would fail to converge.

Regarding the second potential problem, it may be true that one can predict the behavior of a familiar other—a spouse, say—without inferring a trait first. But that is not the issue here. Rather, the concern for accuracy research is how to evaluate a personality judgment once it is made. A judgment that can predict behavior is more likely to be accurate, all other things being equal, than one that cannot, regardless of whether the same prediction could also be made in other ways.

Ordinarily, the combination of consensus and behavioral information takes the form of convergence or, in more colloquial terms, the "duck test" (Block, 1993). Something that looks, sounds, and acts like a duck may not necessarily be a duck, but it very probably is. Similarly, when a person reports that he or she is sociable, is described by several acquaintances as sociable, is quick to initiate conversation with a stranger in a laboratory situation, and spends most of his or her spare time in social activities, then he or she very probably is sociable, no matter how elusive such a determination may be as a matter of ultimate philosophical truth. So when someone describes this person as sociable, it is reasonable to regard this judgment as accurate. And when someone describes this person as unsociable, it is reasonable to regard this judgment as wrong.

In more complex cases, discrepancies among data sources may be as informative as their convergence (McCrae, 1994; Ozer, 1989). For example, people generally give themselves higher ratings than do observers on "internal" traits like "concerned with philosophical problems," whereas observers give higher ratings on "external" traits like "initiates humor." At the same time, self-ratings and observer ratings manifest good correlational agreement on most of these same traits (Funder, 1980a). This set of findings shows how the differing perspectives of self and other change the salience of the thoughts and behaviors associated with internal and external traits. In another example, narcissists generally rate their contributions to a group as more valuable than do expert observers. This is a discrepancy that only underlines the accuracy of their characterization as narcissistic (John & Robins, 1994a).

All of these considerations make the evaluation of accuracy seem complicated, which it is. A realistic approach puts exceptionally heavy conceptual and empirical burdens on any prospective investigator. And, given ordinary limitations on research resources, no one can hope to gather all the criteria one might desire in any particular study. But the goal set by RAM is always to gather as many criteria as you can and for the literature as a whole not to restrict oneself to just one or a few criteria for accuracy. Perhaps ironically, the fact that all potential criteria for accuracy have fundamental limitations means that research must use all—not one or a few—of them. As the accuracy literature continues to develop, the prescription of RAM is that the criteria used should become increasingly diverse.

Accuracy projects. An example of the kind of multifaceted investigation dictated by these considerations is the research by Gifford and his associates on the encoding and decoding of observed behavior. In several painstaking studies, Gifford has shown how personality affects both nonverbal and verbal behavior that, in turn, influences observers' perceptions of personality that, as a result, achieve a reasonable degree of accuracy (Gifford, 1994; Gifford & Hine, 1994).

Another example is the Accuracy Project conducted for the past decade by myself and my colleagues (Funder, 1989,

1993a). This project has assembled two large data sets. The first set includes self and peers' judgments of personalities of about 160 Harvard undergraduates, along with videotapes of their behavior in three laboratory situations. These data have revealed good self–other and interjudge agreement on most attributes of personality and impressive predictive validity between these same personality judgments by self and peers and the videotaped behavior samples (Colvin et al., 1995; Funder & Colvin, 1988, 1991; Funder & Dobroth, 1987; Funder & Sneed, 1993; Kolar, Funder, & Colvin, in press).

An even larger data set has recently been completed at the University of California, Riverside. These newer data include personality judgments by the self, college peers, hometown peers and parents, seven videotaped behavior samples, and 8 days of diary and experience sampling ("beeper") data reflecting the subjects' behavior and emotions in daily life. Similar findings of convergence and predictive validity are beginning to emerge from these newer data (Funder et al., 1995). The overall goal of the Accuracy Project is to extend the domain of data that can serve as criteria for the accuracy of personality judgment as far as is possible, but of course much more remains to be done, in our lab and elsewhere.

The Path to Accurate Judgment

RAM's realistic assumption implies that a personality trait can be accurately judged if the judge can manage to detect and correctly use behaviors that are relevant to the trait and available to his or her observation. To rearrange these steps into their sequential order, the path to accuracy leads through four aspects of the process of personality judgment. First, in some context or contexts, a trait produces a behavioral effect. The resulting behavior is then relevant to that trait. Second, this behavior must be available to the judge. A behavior that is a result of an action of the hypothalamus is not ordinarily available to an observer; neither is a behavior that happens in a context in which the observer is not present. Third, the relevant and available behavior must be detected by the judge. This is a hazardous step, because the judge may be inattentive, unperceptive, or distracted, or the behavior itself (e.g., a momentary facial expression), although available in principle, might be extremely difficult to see. Fourth, the relevant, available, detected behavior must be correctly utilized by the judge. This step is hazardous as well, because the judge might believe a behavior to be diagnostic of one trait when it really is diagnostic of a different one, or of nothing at all.[5]

One way to represent this pathway is shown in Figure 1. The figure, which is a descendant of the "lens model" pioneered by Brunswik (1956), illustrates the steps that lead from attributes of the target person, through his or her behavior, through observation and evaluation of that behavior, to a possibly accurate judgment.

Figure 1 describes a process that contrasts sharply with those posited by more cognitively oriented approaches such as the error paradigm. Much of RAM's process occurs in the social environment along the path between reality and its perception. In contrast, cognitive approaches have focused on the last stage depicted, in which information already detected is used in a more-or-less normative fashion within the mind of each individual perceiver. The perceivers neglect the first two stages, in which a trait manifests itself in behavior, and have paid relatively little attention to the third stage, in which the information is or is not detected.

Integrating personality and social psychology. Figure 1 further illustrates how RAM repeatedly crosses the traditionally well-defended border between personality and social psychology (Funder & Sneed, 1993). Social psychology has concerned itself with how cues are detected and used in judgment without addressing the degree to which these cues are actually diagnostic of the traits they are used to judge under realistic circumstances. Its approach has been almost exclusively experimental, in which cues are manipulated and serve as the independent variable and judgments serve as the dependent variable.

In a parallel fashion, personality psychology has concerned itself with the behaviors that tend to be emitted by people who are characterized by certain traits, but has not addressed the manner in which laypersons might use these behaviors when inferring these traits. Its approach has been almost exclusively correlational, in which subjects' preexisting personality traits serve as the independent variable and their behavior serves as the dependent variable.

Notice how the dependent variable of personality psychology becomes the independent variable of social psychology in RAM's conceptualization. Thus, the left half of Figure 1 can be seen as representing the traditional domain of personality psychology, and the right half can be seen as representing social psychology's typical approach to person perception.

RAM's formulaic representation. An alternative representation of RAM is in terms of the following formula:

accuracy

= [(the relevance of behavioral cues to a personality trait)

× (the extent to which these cues

are available to observation)]

× [(the extent to which these cues are detected)

× (the way in which these cues are used)]

The formula emphasizes different aspects of RAM than does Figure 1. For example, the brackets separate the "environment" side of this equation from the "perceiver" side, but the formula's multiplicativity implies that the distinction between perceiver and environment may be less conceptually sharp than is sometimes assumed. They are both part of the same interactive system, an insight that only further blurs the traditional separation between personality and social psychology.

---

[5] A further complication is introduced by the method through which the judgment is operationalized. The judge might be asked to report his or her judgment on a questionnaire, through a Q-sort, or in a free-response format. Each of these methods introduces its own possibilities for and obstacles to accuracy. More general influences on reporting of judgments, such as response sets, also might come into play. These influences on the reporting of judgments, however, as opposed to the making of judgments, lie outside of the RAM model at present.

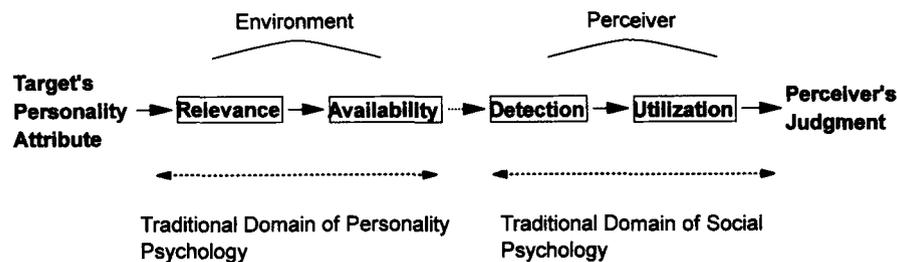## A Model of the Process of Accurate Personality Judgment

Figure 1.   A model of the process of accurate personality judgment.

The multiplicativity also highlights how some degree of availability of relevant cues and some degree of correct observation and inference are required for accurate judgment. If any term in this formula equals zero, then so does accuracy. All the relevant cues in the world are no help if the judge does not perceive and use them; even the most astute judge is helpless in the face of a lack of relevant cues.

A further implication of the formula is that accuracy is, at best, a probabilistic matter. Perfect accuracy can be attained only when all terms in the equation equal one, representing perfectly unambiguous and visible cues to the judgment together with optimal observation and integration of those cues. These kinds of perfection are all theoretical limits rather than empirical likelihoods; therefore, perfection of judgmental outcome represents such a limit as well.

The formulaic representation of RAM also helps to explain why accurate judgment is so difficult. Because the fractions of each term are multiplied, sources of error combine multiplicatively as well. For example, if each term were to represent 90% fidelity, then overall accuracy would be .66; if each term were to represent 50% fidelity, then total accuracy would only be .06. In fact, and in a manner similar to other formulaic representations of social judgment theory (cf. Kenny, 1991, 1994), there is no developed technique for assigning actual numbers to the terms in the RAM equation. Still, the formula does imply that only when all four links in the process of accurate judgment are strong can any substantial degree of accuracy be anticipated. Because good accuracy is a widely obtained outcome, a further implication is that in general each of these links—relevance, availability, detection, and utilization—may have a rather high value under many circumstances.

Multiple cues and multiple traits.   The RAM model is a simplified description of a prototypic part of the way personality judgment happens in real life. The model essentially describes a one cue–one trait process in which a single piece of information (e.g., a behavior) is or is not relevant, available, detected, and used on the way to the accurate judgment of a single trait. Social reality is of course more complicated than that. In a typical case, it seems likely that people simultaneously, or nearly simul-

taneously, detect and use numerous cues toward the judgment of numerous traits (Funder & Sneed, 1993).

Moreover, this complexity might be more than common; it might be important. It seems highly possible that the interactions among nearly simultaneous judgments of different traits might affect how each trait is judged. An environment that contains multiple cues certainly will yield judgments affected by interactions among those cues as well as by each cue considered alone (Borkenau & Liebler, 1992, 1993a). In addition, many different cues might be diagnostic of the same trait, whereas the same cues might be simultaneously diagnostic of different traits (Buss & Craik, 1983; Funder, 1991).

One way—a shortcut—out of this tangle of complexity would be to broaden the interpretation of the RAM model. It can be noted that, at a general level, the cues described in Figure 1 are multiple, as are the traits judged, and these cues and judgments interact as well as flow from one to the other. This point could simply be acknowledged, or Figure 1 could be redrawn, with numerous lines running from left to right and numerous two-headed arrows interconnecting all of the lines.

Such a re-representation of RAM would more properly acknowledge the complexity of the phenomenon it seeks to model and would even be rather humbling, to be sure. But such a convolution would accomplish little else at this time. To understand how cues and judgments interact will require more than acknowledgment that they probably do. It will require research, in which simultaneous cues are measured and simultaneous judgments detected, to describe what specific cues or judgments interact with what other specific cues or judgments, and how and why. RAM offers an initial, prototypic model sufficient to account for much of what is now known about accurate personality judgment and to organize research that is likely to occur in the near future. An appropriate task for the next generation of research will be to track down and empirically demonstrate specific interactions among cues and judgments and their effects on accuracy.

Implications for Moderators

One of the goals of RAM is to provide a common means for accuracy research by accounting for the diverse variables that

Table 1
Aspects of Judgmental Process Associated With Moderators of Accuracy

Moderator	Specific characteristic	Relevant RAM process variables
Good judge	Perceptiveness	Detection
	Judgmental ability	Utilization
	(Non)defensiveness	Detection, utilization
Good target	Activity level	Availability
	Consistency, scalability	Relevance
	Ingenuousness	Availability, relevance
Good trait	Visibility, frequency	Availability
	Operant/respondent	Relevance
	(Non)evaluativeness	Availability, relevance
Good information	Quantity (e.g., acquaintance)	Availability
	Quality (e.g., relationship)	Relevance

Note. These are representative examples of the process variables RAM proposes underlie various specific moderators of accuracy, organized by four broad categories. RAM = Realistic Accuracy Model.

affect accuracy in terms of a common process model. RAM does this by assuming that any variable that makes accuracy more or less likely must gain its effect by virtue of its relevance to one or more of the terms in the equation mentioned earlier. It must involve relevance, availability, detection, utilization, or more than one or all of these. Table 1 lists the four basic moderators of accuracy and some examples of each, along with the relevant stage or stages in the RAM model depicted in Figure 1.

Good judge. Differences between judges of personality who are generally accurate and those who are generally inaccurate must, according to RAM, be produced by their differential detection or the use of available cues. The ability to perceive and to use available cues correctly can be divided into three components. The first is knowledge about personality and how it is revealed in behavior, which will ordinarily be a function of the amount and type of the judge's interpersonal experience, but might under some circumstances stem from explicit teaching or study. Such knowledge might enable a judge to detect behavioral cues others could miss. This hypothesis is consistent with the conclusion by Akert and Panter (1988) that "because extraverts have more experience in social settings than introverts" (p. 965) they are better at decoding nonverbal cues in social interaction. Extraverts also seem to be better at distinguishing real from simulated suicide notes (Lester, 1991).

A second component is ability, which might derive from broad variables such as IQ as well as more narrow abilities such as cognitive or attributional complexity. Such judgmental ability might improve the chances of such cues as are detected to be used in a valid manner. At least two studies have found that judges with higher measured intelligence more accurately rated the performance and emotions of others (Havenstein & Alexander, 1991; Westbrook, 1974).

A third component is motivational, and may take two forms. The first is simply the degree to which the judge cares about the judgment he or she is making and whether it is accurate. Flink and Park (1991) found that when subjects were told that important social outcomes depended on the accuracy of their judgments, their judgments were made with greater consensus.

Moreover, a judgment that one does not care to make in the first place can scarcely be accurate.

Another aspect of motivation that is relevant in this context, and illustrated in Table 1, concerns the possibility of defensiveness or other motivated styles of information processing that have the effect of distorting one's perception and judgment. An individual whose thinking is motivated by the intense need to believe himself always to be in the right cannot be expected to have a high degree of accuracy when judging the actions and personalities of other people. Defensiveness and lack of humor about one's own shortcomings should be particularly associated with low accuracy, as should a narcissistic view of the self (John & Robins, 1994a) or generally hostile or exceptionally benign attitudes about others.

Although RAM implies that all the attributes just discussed may be associated with judgmental accuracy, it does not imply that self-assessments of ability will necessarily predict accuracy. Whereas a good judge is likely to know that he or she is good, there is little stopping a poor judge from also believing himself or herself to be a good judge. His or her very lack of insight may create an assessment conundrum, making it difficult to detect this lack of insight through self-report (Fletcher, 1993; Ickes, 1993).

As mentioned earlier, although historically the good judge is the first potential moderator to have been addressed by research, it remains the one for which, to date, the accuracy literature has the sparsest data and fewest firm findings to report. One reason may be that early studies were plagued by methodological and analytic shortcomings, of the sort identified by Cronbach (1955), that have only recently begun to be rectified (Bernieri, Zuckerman, Koestner, & Rosenthal, 1994; Funder & West, 1993a). Another reason may be that extensive data about the personalities or abilities of the informants who provide personality judgments have been obtained only rarely. Despite the difficulties encountered by accuracy research in the 1950s, recent evidence does indicate that at least some kinds of judgmental ability are general across targets (Borman, 1977; Marangoni, Garcia, Ickes, & Teng, 1995; but see Gangestad, Simp-

son, DiGeronimo, & Beik, 1992). So it is still premature to give up on the good judge, and RAM provides several hypotheses about what the concomitants of judgmental ability might prove to be.

Good target. The personality and general behavioral patterns of some individuals can be judged correctly from relatively few observations of their behavior, whereas others remain an enigma even after prolonged observation (Allport, 1937; Colvin, 1993a, 1993b). Allport's (1937) classic question in this context was, "Who are these people?" (p. 443). According to RAM, individual differences in the tendency to be judged accurately are a matter of cue availability and relevance. People whose behaviors comprise numerous and informative (e.g., relevant, nonmisleading) clues to their personalities should be the easiest to judge.

This simple principle yields a number of hypotheses. For example, some people have more behaviors than do others as a function of their general activity level. Others, those who are relatively inert, do less. RAM predicts that, all other things being equal, people with a higher level of behavioral activity, especially social activity, should be easier to judge than those who are less active, because more active people naturally have more behaviors that then become available as potential clues about their personality (e.g., Borkenau & Liebler, 1992). People who do not do or say as much, in contrast, will be more difficult to judge (beyond, perhaps, the judgment that they are passive or withdrawn).

Turning to more complex phenomena, self-monitoring theory (Snyder, 1987) describes individuals who adjust their behavior with great sensitivity to even subtle changes in the surrounding social environment. These individuals, called *high self-monitors,* should be more difficult to judge accurately than *low self-monitors,* who are theorized to more likely be themselves and to act consistently across diverse situations. In a related vein, Bem and Allen (1974) discussed differences between people who were and were not consistent in their behavior across situations and showed that those people who described themselves as consistent on a given trait were judged on that trait with better agreement by others. Although the initially demonstrated effect has not always been successfully replicated by others (Chaplin & Goldberg, 1984), the central claim that individual differences in predictability can be identified has held up reasonably well (Zuckerman, Bernieri, Koestner, & Rosenthal, 1989).

Bem and Allen's (1974) basic ideas have evolved over the years in the hands of other investigators, yielding Baumeister and Tice's (1988) notion of "metatraits" (the trait of having a trait), as well as Lanning's (1988) notion of "scalability." The idea of metatraits seeks to distinguish between people who are and are not consistent in their behavior along a given trait dimension (see also Tellegen, 1988). It harkens back to Bem and Allen's classic demonstration of how more and less predictable people could be distinguished from each other simply by asking them, "How consistent are you?"

The notion of scalability, also discussed by Bem and Allen (1974), is that the behavior of some individuals is not patterned as ordinary trait constructs would predict. In psychometric terms, these individuals pass the "hard items" but flunk the "easy items." Bem and Allen pointed out that in general people find it easier to be relaxed and friendly one-on-one than in front of a group of 400 observers (such as a class full of students), and that is the usual structure of sociability. But some individuals—such as Daryl Bem himself—find it easier to behave sociably on a public stage than in a private setting, and therefore are not scalable in the ordinary way. Reise and Waller (1993) have pushed this idea in a more specifically psychometric direction, developing the measurement of scalability using item response theory and assessing differences between individuals who are and are not scalable on particular traits as well as in general. RAM would predict that people should be particularly difficult to judge along the dimensions for which they are "untraited" or unscalable.

Colvin's (1993a, 1993b) conception of the "judgable" person, which grew out of the present research program, is an effort to integrate differences of the sort just discussed. Some individuals, according to Colvin's analyses, are both more consistent in their behavior across experimental situations and more likely to be agreed about by diverse informants. This is not because their judgability is a trait in itself; rather, judgability is an entailed manifestation of behavioral consistency. A person whose words, deeds, and thoughts are all consistent with each other at every moment and across situations becomes extremely easy to judge, by definition. A word can be used to predict a deed, a deed to predict a thought, and an action at one time to predict an action at another time all because the behavior of this person is consistent and coherently organized. In this light it becomes unsurprising that judgability is associated with sound psychological adjustment and mental health; it is the natural concomitant of a well-organized personality and consistent behavioral style (Colvin, 1993a, 1993b; Jourard, 1971).[6]

A nonjudgable person, by contrast, is one whose actions and thoughts cannot be predicted from observations of what he or she says or does. This might be because he or she is concealing something, is self-aggrandizing, or is high in self-monitoring, or even more basically because he or she has an incoherent and disorganized personality—perhaps, in some cases, to the point of mental illness. This conclusion fits well with the results of Donahue, Robins, Roberts, and John (1993), who reported evidence that being inconsistent, or "seeing oneself as having different personality characteristics in different social roles . . . is a sign of fragmentation of the self" (p. 834). It is also consistent with the recent findings of Reise and Waller (1993), who reported that being less scalable in general, across traits, is associated with lower well-being, less adaptive reactions to stress, more alienation, more aggression, and less self-control.

A final implication is that individuals motivated to conceal aspects of their personalities will be more difficult to judge. Thus, people with dishonest or otherwise socially undesirable tendencies will generally seek to conceal them, leading themselves to be difficult to judge accurately on the basis of their overt social behavior (Aronson & Mettee, 1968; Kuiken, 1981). Individuals with more socially acceptable tendencies have no

---

[6] An important further research question eventually becomes, At what point, or under what circumstances, does adaptive consistency cross the line into maladaptive rigidity?

reason to conceal them, by contrast, and so may be more likely to display behaviors that are indicative of their true selves.

Good trait. Judgability can be a property of traits as well as of individuals. From the perspective of RAM, the difference between traits that are more and less easy to judge derives from the existence of cues to their judgment that are available and relevant.

Concerning availability, a trait like sociability, which is revealed by frequent positive social interaction, is easier to judge than a trait like ruminates and daydreams, which must be inferred from verbal statements or, even more ambiguously, from dreamy looks, distracted responses, and the like, any of which could have other meanings as well or instead. Research shows that visibility, in this sense, is closely associated with interjudge agreement. This basic finding was reported by Funder and Dobroth (1987) concerning agreement between the self and knowledgable informants and was extended by Funder and Colvin (1988) to agreement among judgments by the self, knowledgable informants, and strangers. Consistent results also have been reported by Bernieri et al. (1994), Borkenau and Liebler (1992), Kenrick and Stringfield (1980), Levesque and Kenny (1993), Watson (1989), and others.

Availability may ordinarily be related to relevance but, as the RAM model illustrates, the two can be separate. For example, talkativeness is a highly visible, available behavioral cue but is also ambiguous because it might be relevant to sociability, nervousness, dominance, or a complicated combination of all of these traits. By the same token, the act of saving a family from a burning building might be unambiguously relevant to the trait of courage but is not ordinarily available because opportunities for the act are so rare. The act of stealing might be unambiguously relevant to the trait of dishonesty but is not ordinarily available because it is typically concealed by the actor as much as possible (see Rothbart & Park, 1986).

Traits also differ from one another in their evaluative or connotative meaning. Some traits are those all people would wish to possess and display (such as courage and intelligence); others, if people had them, they might prefer to avoid displaying (such as dishonesty). The motivations for self-presentation that naturally result from this difference can be expected to make some traits less often or less directly displayed—or exaggerated—in overt behavior, leading to distortions in their availability and relevance. As a result, they would tend to be judged less accurately. In support of this hypothesis, John and Robins (1994b) reported that traits that are extremely desirable or undesirable tend to yield lower self–other agreement in their ratings, compared with more neutral traits. John and Robins inferred that judgments of evaluatively loaded traits are prone to be distorted by self-protective and self-enhancing motivational processes.

Good information. This moderator of accuracy refers to the information that is available to the judge independently of the degree that he or she might perceive it or the manner in which he or she might use it. For personality judgment, information consists of anything the person who is judged says or does that might be relevant to the kind of person he or she is.

A simple but important consideration is whether the judge has observed the target person enough to enjoy the reasonable possibility of making an accurate judgment. Recent research

has shown that acquaintances who have known their targets for about a year agree with each other and with targets' self-judgments better than do relative strangers who have viewed the targets only by means of a single 5-min videotaped behavioral episode (Funder & Colvin, 1988), a finding parallel to that obtained by other investigators (Cloyd, 1977; Colvin & Funder, 1991; Funder & Colvin, 1988; Jackson, Neill, & Bevan, 1969; Norman & Goldberg, 1966; Paulhus & Bruce, 1992; Paunonen, 1989; Taft, 1966).

The most parsimonious explanation of this finding is that more information is available to acquaintances than to strangers, making their judgments more accurate. Recent research has supported this interpretation by eliminating possible alternative explanations such as "assumed similarity" and discussion among judges (Borkenau & Liebler, 1993b). One recent pair of studies showed that well-acquainted college judges were not more similar to their targets than they were to nontargets in the same college population, nor did their similarity drive their accuracy. Moreover, judges who had met each other agreed in their judgments of the personalities of a third person no better than judges who had not met. These findings indicate that neither assumed similarity nor conversation is necessary for interjudge agreement (Funder et al., 1995; see also Fuhrman & Funder, in press).

Kenny's (1991, 1994) Weighted Average Model (WAM) of the determinants of interjudge agreement also casts quantity of information in a prominent role. His model assumes that the amount of information that is shared by two judges—the observations of the target person they have in common—will affect the degree to which they reach consensus in their judgments. The more there is of such information overlap the more likely judges are to reach consensus.

A problem with the quantity parameter, however, is that the variable is rather crude, certainly as operationalized to date and perhaps even in principle. WAM, for example, offers little guidance about how to measure amount of information or how to define and count behavioral episodes, all of which are essential steps toward testing its predictions. Likewise, because Funder and Colvin (1988) compared accuracy only between judges near either end of the information–acquaintanceship dimension—some knew their targets extremely well, the others knew their targets hardly at all—the issue of how to quantify information was essentially finessed.

Ample evidence indicates that consensus increases with acquaintanceship. But to describe the exact slope and shape of this function will require a well-controlled longitudinal study of the course of acquaintanceship over a significant period beginning with first meeting (see Paulhus & Bruce, 1992, for a pioneering demonstration of how such a longitudinal study can be done). Alternatively, an experimental study could directly manipulate the amount of information available to each judge.

Beyond the sheer amount of information available to a judge, the quality of information also matters. Certain kinds of information may be generally more or less diagnostic of personality. Andersen (1984) demonstrated that listening to a person talk about his or her thoughts and feelings leads to more accurate personality judgment than does listening to that same person talk about his or her activities and hobbies. In a similar vein,

Table 2
Interactions Among Moderators of Accuracy in Personality Judgment

Moderator	Judge	Trait	Target	Information
Judge	—	Expertise	Relationship	Sensitivity
Trait	—	—	Palpability	Diagnosticity
Target	—	—	—	Divulgence
Information	—	—	—	—

Snyder and Ickes (1985) theorized some years ago that observing somebody in an unstructured situation, in which behavior is relatively free to vary, is more generally informative about personality than observing him or her in a situation with more structure and less allowable behavioral variation. This suggestion has been borne out by more recent research, which found that unstructured dyadic interactions yielded more and larger correlations between personality and behavior than did a more structured debate situation (Funder & Colvin, 1991).

The relationship between information and trait relevance can also be quite specific. Scherer (1978), for example, has shown that the simple cue of speaking in a loud voice is a valid indicator of extraversion. Gifford and Hine (1994) have demonstrated links between extraversion and aloofness and observable, verbal behaviors that perceivers in turn validly use to judge these traits. Gifford (1994) found similar linkages with nonverbal behavior.

Interactions Among Moderators

The variables of good judge, good target, good trait, and good information inevitably overlap and interact. For example, a good target is one who emits good information, and certain traits may be more visible in some targets than in others. The four general moderators of accuracy yield six unique interactions, which are shown in Table 2 along with the term by which each is denoted in RAM. These terms each have a technical meaning that is related but not equivalent to their everyday meaning; each term also should be considered provisional and may change as the theory continues to develop.

Judge × Trait: Expertise. Certain individuals might be good at judging some traits but poor at judging others. This condition might arise because of variations in their knowledge across traits or differential ego involvement concerning certain traits. Differential knowledge might be a function of an individual's life experience or explicit teaching and study and might be associated with the differential cognitive availability of certain traits that make information about them more likely to be perceived and used accurately. For example, Park and Judd (1989) found that judges who showed a readiness to judge others' intelligence, honesty, and conscientiousness also tended to judge these traits with greater consensus. In a motivational vein, someone concerned about his or her adequacy in a certain domain might be poor at judging the adequacy of others in that same domain. Self-serving differences in judges' conceptions of traits can be an important source of interjudge disagreement and, presumably, of inaccuracy (Dunning, Perie, & Story, 1991).

The particular interactive relationship between a specific judge and his or her ability to judge accurately a specific trait, as opposed to traits in general (which is the main effect of judge), is called *expertise* within RAM. This term refers not just to knowledge or skill but to a lack of emotionally relevant or motivated distortions in processing certain kinds of information. Therefore, RAM's theoretical usage of the term expertise does not exactly match its meaning in lay language. Expertise is relevant to the utilization stage in RAM, in which information already relevant, available, and detected is or is not evaluated and combined in a manner that produces an accurate judgment.

Judge × Target: Relationship. Certain judges might be able to judge some targets more accurately than they can judge others, who might in turn be more accurately judged by still other judges. This interaction could occur when a target's most central traits happen to be those about which a particular judge has expertise, as just defined, or when the relationship between judge and target serves to enhance rather than interfere with accuracy. For example, a competitive relationship between judge and target, because it interferes with accurate cue utilization, can be expected to diminish accuracy. In a similar vein, Sillars and Scott (1983) have observed that although marriage and dating partners might be extremely knowledgeable about each other, the nature of their relationship and their interdependency can make objectivity in their mutual perceptions difficult to attain (see also Stiff, Kim, & Ramesh, 1992). For example, dating couples in insecure relationships may avoid detecting the degree to which their partner is attracted to another person (Simpson, Ickes, & Blackstone, 1994).

Still another possible basis for this interaction is that a judge might evoke informative behaviors from a particular target, perhaps because of their mutual ego involvement, that would not be visible to another judge lacking such evocative effect. For example, colleagues, lovers, people who are attracted to each other, and people who loathe each other probably all display behaviors in each others' presence that are not visible to other observers and therefore may be able to judge each other on dimensions that others could not. (To the degree this effect is trait specific, it could be denoted by a three-way Judge × Target × Trait interaction.)

Another aspect of this interaction can be considered separately. That is the probable fact that one gets different information during the course of acquaintanceship as a function of the setting and type of relationship (e.g., as a coworkers as opposed to as a sibling). Or, as one's relationship develops from acquaintance to friend to lover (and, in some cases, perhaps to enemy), one finds oneself in a position to observe different kinds of be-

havior and thus draw different inferences about the person. This effect of type of relationship on the accuracy of personality judgment has received scant attention in prior research but deserves increasing attention in the future.

Another prediction derivable from RAM concerns gender differences. It has widely been observed that a principal activity of female–female friendship pairs is the discussion of emotions and relationships, whereas male–male pairs are more likely to spend their time in hobby or work activities or in the discussion of more impersonal matters such as sports or politics (Reisman, 1990; Sherrod, 1989). Combining this observation with Andersen's (1984) findings, that conversations that reveal more personal information produce better information on which to base personality judgments, leads to the prediction that female–female information–target pairs should judge each other more accurately than male–male pairs. This particular prediction does not yet seem to have been empirically tested, although the general (albeit small) superiority of women over men in the detection of emotional states is a long-standing staple of the literature (Hall, 1990; Kirouac & Dore, 1985; but see Ambady & Rosenthal, 1992).

RAM's prediction concerning dyadic gender differences is based on the information derived from the distinctive interactional styles of each gender and the information yielded therefrom. It does not stem from any intrinsic quality of men or women. RAM would also predict that a male–male pair that does discuss emotions would be as mutually accurate as any female–female pair and that a female–female pair that engages solely in work or hobby pursuits could be just as mutually inaccurate as any male–male pair. More generally, the basis of all of RAM's predictions concerning the effect of type of relationship on accuracy is the kind of information that is yielded by the behaviors that become visible in it.

Any interaction between two people constitutes a relationship. Therefore, the two-way interaction between judge and target that affects accuracy for any of the reasons just summarized is called the *relationship* variable within RAM. This term is meant to capture a unique alignment between two individuals that can enhance or lessen the accuracy of the judgments that one makes of the other. The term is not equivalent to its meaning in lay usage, which is broader and includes aspects of relationships that might not affect accuracy.

Judge × Information: Sensitivity. Certain judges might prefer or be able to perceive and use certain kinds of information but not other kinds. A judge might be acutely aware of anything that suggests competitiveness or dominance, for example, or might always emphasize any information that becomes available about a target's religious or political leanings. Sedikides and Skowronski (1993) showed that when honesty or intelligence was an important part of a person's self-concept, he or she tended to be particularly attentive to information about the honesty or intelligence of others. Judges also vary in their ability to detect and recognize nonverbal behaviors that may be diagnostic of emotion or personality (Hall, 1990). This tendency of certain judges to search for and perceive certain information, or to weigh certain kinds of information heavily in their judgments, is called *sensitivity.*

In RAM, such sensitivity is regarded as specific rather than general; a judge could be highly sensitive to one kind of information at the same time he or she is oblivious to another. Bargh and Pratto (1986) have shown how some individuals are particularly affected by information relevant to constructs that are "chronically accessible" within their cognitive systems. Similarly, Markus, Smith, and Moreland (1985) reported that the nature of information structures or "schemas" in the self-concept can affect how one perceives and evaluates the behavior of others.

When the information to which a certain judge is sensitive tends to be accurately diagnostic of a particular trait, this interaction becomes equivalent to the Judge × Trait interaction described earlier. But notice how sensitivity is not necessarily expertise. If one is sensitive to information that turns out to be misleading, then such sensitivity would harm rather than enhance accuracy. As noted earlier, expertise is relevant to the utilization stage within RAM. Sensitivity, by contrast, is relevant to the detection stage.

Trait × Target: Palpability. Certain traits might be easy to judge in some targets but not others, or certain targets might have traits that can be judged easily and others that cannot. As in all of these interactions, it is important to keep this interaction separate from the main effects of its two components. The interaction refers to traits that might stand out in certain targets, relative either to the same trait in other targets or other traits in the same target. For example, a particular person's deeply ruminative style might be the most salient and easily judged aspect of him or her, even though in general this is one of the least visible traits (Funder & Dobroth, 1987; Kenrick & Stringfield, 1980). Or, a person's high degree of anxiety might be clearly visible even while his or her other traits are almost completely obscure.

Research by Bem and Allen (1974) as well as more recent work on traitedness by Baumeister and Tice (1988) and on scalability by Lanning (1988) and by Reise and Waller (1993) usually has focused on individual differences in the consistency of particular traits. Bem and Allen looked at differential consistency in the traits of friendliness and conscientiousness, and later work too has usually assessed traitedness or scalability one or two traits at a time (Koestner, Bernieri, & Zuckerman, 1989). Therefore, most of this earlier work can be seen as more directly addressed to this Trait × Target interaction than to the main effect of target discussed earlier.

Traits that are central to a person's self-concept tend to be easier for others to detect, leading to greater interjudge agreement (Koestner, Bernieri, & Zuckerman, 1994). The reason seems to be that people are motivated to be seen by others in ways that verify their self-concepts, and certain central traits are not only important to the self-concept, but serve to organize it (Sedikides, 1993). RAM calls the property of certain traits to be particularly judgable in certain persons *palpability,* in the sense of being obvious and therefore judgable, against other traits that are less so.

Trait × Information: Diagnosticity. Some traits can be judged only on the basis of particular kinds of information. For example, Colvin and Funder (1991) showed that viewing a 5-min videotape was a relatively poor (although far from useless) source of information for judgments of general aspects of per-

sonality but was as good a source for predictions of the target's future performance in a specific, similar situation as was that derived from long acquaintanceship. In a related vein, Paunonen (1989) showed that whereas some traits can be judged accurately (by the criterion of self–other agreement) on the basis of quite minimal observation, others can be judged accurately only on the basis of extended acquaintanceship.

This kind of specific relation between the information available and the particular trait judged in RAM is called *diagnosticity*. The sense is that certain kinds of information lead relatively directly to the accurate judgment or diagnosis of certain traits, whereas other information may be irrelevant, insufficient, useless, or even misleading.

Target × Information: Divulgence. Certain kinds of information might tell a judge a great deal about one target but relatively little about another. A particular individual's impoverished upbringing, for example, might go a long way toward explaining a wide range of his or her actions and traits. The same background in someone else might signify relatively little. Similarly, one individual's racial or ethnic identity might be a key to understanding his or her personality, whereas for another member of the same group his or her ethnicity might have little or nothing to do with the kind of person he or she has become (Azibo, 1991).

Information that might in this sense reveal all is described in RAM by the term *divulgence*. It is meant to imply that all information about a given individual is not created equal and certain kinds might disclose and reveal (the dictionary definition of *divulge*) much of the personality of one individual whereas similar information reveals little about another.

Self-Judgments

As introduced in this article, RAM addresses only processes that arise when one individual judges the personality of another. An important further question, just beginning to be examined, concerns the extent to which the process of accurately judging others extends also to judgments of and insight into oneself (Hofstee, 1994; Robins & John, in press). In the past, some social psychologists have treated self-perception as essentially a special case of the process of other perception (Bem, 1972). Others have emphasized what they see as fundamental differences between judgments of the self and of others (e.g., Greenwald, 1980; Jones & Nisbett, 1971). An important future puzzle for RAM will be to determine whether, when, and how its key concepts—relevance, availability, detection, and utilization—can be extended to the case for which the target of judgment is oneself.

For example, in a recent article, Kolar et al. (in press) compared the predictive accuracy of personality judgments by the self and by close acquaintances. They found that taken singly these two sources yielded judgments that demonstrated approximately equal validity for predicting independent, direct observations of the targets' behavior in the laboratory. But the average of as few as two others' judgments significantly outperformed self-judgments.

Kolar et al. (in press) speculate that the natural and obvious informational advantages of the self might be balanced by other factors that favor outside observers. One possibility is that people distort some of their self-observations to protect their self-esteem, making their self-judgments less accurate (Colvin et al., 1995; John & Robins, 1994a). Such an effect would occur at the utilization stage of RAM. Another possibility is what Kolar et al. (in press) called the *fish and water effect*. This term refers to the possibility that our own behavioral patterns are so familiar to us that we become relatively unable to perceive them, for roughly the same reason that fish are said to find it difficult to perceive water. Accurate judgment of personality might then become more difficult for the individual himself or herself than it would be for an outside observer (see also Hofstee, 1994). This effect would occur at the detection stage of RAM.

Discussion

SRM, WAM, and RAM

RAM has some important points of contact with as well as divergence from Kenny's (1991, 1994) Social Relations Model (SRM) and Weighted Average Model (WAM) of interpersonal perception.

The SRM provides a method for collecting and analyzing data using a round-robin design in which all judges make judgments of all targets and vice versa. Data analysis decomposes the variance of judgments into analysis of variance-like components attributable to the target, trait, and judge and their interactions. Such decomposition is useful for addressing nine basic questions about interpersonal perception, one of which is target accuracy, or "Is [the perceiver's] view of [the target] correct?" (Kenny, 1994, p. 6).

RAM shares the SRM's emphasis on the study of real people who have had a chance to observe or to interact with each other. However, RAM and the SRM differ in at least three ways. First, accuracy is a central rather than a peripheral concern of RAM. In the SRM, accuracy is listed as the fifth out of nine issues addressed; the other eight concern the degree to which different kinds of interjudge agreement—independent of accuracy—are present (Kenny, 1994, Table 1.1). Second, studies consistent with RAM do not ordinarily use round-robin designs of the sort dictated by the SRM. Some research questions addressed by RAM involve perceptions among people who have known each other for considerable periods, and some do not require the SRM's often useful but sometimes procedurally prohibitive design (Funder & West, 1993a). Third, as Kenny acknowledges (1994, p. 214), the SRM has difficulty handling moderator variables, and research within the model to date has tended to neglect them. Such moderators are a central concern of RAM.

A further approach, WAM, addresses psychological rather than methodological and data analytic issues (Kenny, 1994). WAM describes in formal terms how the effects of culture, stereotypes, communication, common observation, and personality mesh to determine the degree to which two judges of the same individual agree. This model does much to specify some influences on interpersonal perception that heretofore have been discussed in only the vaguest of terms.

The central point of divergence between WAM and RAM is that they are aimed at distinct phenomena. The critical depen-

dent variable within WAM is interjudge agreement; the model specifies how variables combine to determine individuals' perceptions and, by extension, the degree of agreement among the perceptions of different individuals. In that sense, WAM is essentially constructivist rather than realistic. The critical dependent variable within RAM, by contrast, is accuracy. The purpose of RAM is to show not only how perceptions arise but how accurate judgment can arise through a combination of properties and actions of the target person and observation and utilization of information by the perceiver. In WAM, accuracy is one variable that affects agreement. In RAM, agreement is one indicator of accuracy.

Another Look at Error and Accuracy

The essence of the difference between the error and accuracy paradigms is that they address (and nearly always disconfirm) different null hypotheses. For the past two decades, the implicit null hypothesis underlying error research has been that human judgment is perfect. This hypothesis generates studies that specify the conclusion judges normatively should reach and then documents real subjects' inevitable failure to precisely attain this criterion (N. H. Anderson, 1990). The theoretical task is then to explain how this lack of perfection occurs (e.g., through the use of judgmental heuristics), and the applied task is to develop procedures that make whatever went wrong less likely.

The implicit null hypothesis underlying accuracy research, by contrast, is that human judgment is always wrong. This hypothesis generates studies that compare judgments with one or more criteria for accuracy, such as interjudge agreement or behavioral prediction, and then documents their attainment of some degree of accuracy, so defined. The theoretical task is then to explain how this modicum of accuracy occurred (such as both WAM and RAM try to do), and the applied task is to develop procedures to make whatever went right more likely.

The opposite methodologies of these two basic paradigms provide useful and even essential complements to each other. Each approach dictates different methods, theoretical questions, and applications, all of which are necessary for a complete science of social judgment.

The differing null hypotheses also sometimes have served more to obscure than to clarify matters. Especially in its early days, error research more than once trumpeted the conclusion that human judgment is not perfect as if it were news and worse, as if it implied the performance of the human judge is generally abysmal. Accuracy research still is in its relatively early days, so perhaps it is understandable—if no less unfortunate—that it has fallen, more than once, into an equivalent pitfall. Accuracy research has occasionally trumpeted the conclusion that people are sometimes accurate as if that were news and as if it implied human judgment is nearly infallible.

The problem is that it was (or should have been) obvious from the beginning that human judgment is sometimes right and sometimes wrong. Any further characterization of it as pathetic or admirable is a value judgment that depends more on what one expected to find and perhaps on the degree of one's dispositional optimism or pessimism than on scientific evidence.

The difficult task for the next generation of research shall be to benefit from the contributions of both approaches while eschewing the more problematic tendencies of each. Error research has already provided a useful catalog of many ways in which judgment can go wrong, illuminated judgmental processes or heuristics that underlie error, and suggested ways in which errors might be avoided (Baron, 1988). The aim of this article and of RAM is to help move accuracy research toward parallel contributions. Beyond demonstrating that accurate judgment sometimes occurs, accuracy research also must demonstrate circumstances under which judgment is most likely to go right, illustrate the processes of accurate judgment, and provide its own suggestions for how accuracy might be improved.

The Purposes of RAM

As the pace of research within the error paradigm begins to slacken, research within the accuracy paradigm has accelerated. The main purpose of RAM is to provide a theoretical guide toward organizing the rapidly accumulating findings of this research, by using a relatively small set of process variables to account for diverse moderators of accuracy.

From the perspective of the error paradigm, a judgment is accurate if it was made according to the prescriptions of a normative model. From the perspective of the pragmatic approach to accuracy, a judgment is accurate if it proves useful to the judge. From the perspective of the constructivist approach, a judgment is accurate if a community of judges agrees that it is. But from the perspective of RAM, none of these criteria is enough. Its realistic approach prescribes that the same amount of effort go into gathering and synthesizing diverse information about the target's personality as into examining what goes on within the head of each judge.

A second purpose of RAM, therefore, is to redirect some research attention back to properties of the individual—the stimulus person—who is the target of personality judgment. A personality trait is something that is real, interesting, important, and sometimes hidden. Its judgment, and the evaluation of its judgment, can only be attained through the most circuitous and difficult of routes. A realistic approach to the study of accuracy implies that personality and social psychology need to brave this route together and seek truth as well as shun error.

References

Akert, R. M., & Panter, A. T. (1988). Extraversion and the ability to decode nonverbal communication. *Personality and Individual Differences, 9,* 965–972.

Allport, G. W. (1937). *Personality: A psychological interpretation.* New York: Holt, Rinehart & Winston.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111,* 256–274.

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

Anderson, J. R. (1990). *Cognitive psychology and its implications.* New York: Freeman.

Anderson, N. H. (1990). A cognitive theory of judgment and decision.

In N. H. Anderson (Ed.), *Contributions to information integration theory* (pp. 105–142). Hillsdale, NJ: Erlbaum.

Aronson, E., & Mettee, D. R. (1968). Dishonest behavior as a function of differential levels of induced self-esteem. *Journal of Personality and Social Psychology, 9,* 121–127.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41,* 258–290.

Azibo, D. A. (1991). An empirical test of the fundamental postulates of an African personality metatheory. *Western Journal of Black Studies, 15,* 183–195.

Bargh, J. A., & Pratto, F. (1986). Individual construct accessibility and perceptual selection. *Journal of Experimental Social Psychology, 22,* 293–311.

Baron, J. (1988). *Thinking and deciding.* Cambridge, England: Cambridge University Press.

Baumeister, R. E., & Tice, D. M. (1988). Metatraits. *Journal of Personality, 30,* 571–597.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). New York: Academic Press.

Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review, 81,* 506–520.

Bem, D. J., & Funder, D. C. (1978). Predicting more of the people more of the time: Assessing the personality of situations. *Psychological Review, 85,* 485–501.

Bernardin, H. J., & Pence, E. C. (1980). Effects of rater training: Creating new response sets and decreasing accuracy. *Journal of Applied Psychology, 65,* 60–66.

Bernieri, F. J., Zuckerman, M., Koestner, R., & Rosenthal, R. (1994). Measuring person perception accuracy: Another look at self–other agreement. *Personality and Social Psychology Bulletin, 20,* 367–378.

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., Block, J. H., & Harrington, D. (1974). Some misgivings about the Matching Familiar Figures Test as a measure of reflection-impulsivity. *Developmental Psychology, 10,* 611–632.

Block, J., & Funder, D. C. (1986). Social roles and social perception: Individual differences in attribution and error. *Journal of Personality and Social Psychology, 51,* 1200–1207.

Block, J., & Robins, R. W. (1993). A longitudinal study of consistency and change in self-esteem from early adolescence to early adulthood. *Child Development, 64,* 909–923.

Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology, 62,* 645–657.

Borkenau, P., & Liebler, A. (1993a). Consensus and self–other agreement for trait inferences from minimal information. *Journal of Personality, 61,* 477–496.

Borkenau, P., & Liebler, A. (1993b). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *Journal of Personality and Social Psychology, 65,* 546–553.

Borman, W. C. (1975). Effects of instructions to avoid halo error on reliability and validity of performance ratings. *Journal of Applied Psychology, 62,* 64–69.

Borman, W. C. (1977). Consistency of rating accuracy and rating errors in the judgment of human performance. *Organizational Behavior and Human Performance, 20,* 238–252.

Borman, W. C. (1979). Format and training effects on rating accuracy and rater errors. *Journal of Applied Psychology, 64,* 410–421.

Bruner, J. S., & Tagiuri, R. (1954). The perception of people. In G. Lindzey (Ed.), *Handbook of social psychology* (pp. 634–654). Stanford, CA: Stanford University Press.

Brunswik, E. (1956). *Perception and the representative design of experiments.* Berkeley: University of California Press.

Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review, 90,* 105–126.

Chaplin, W. F., & Goldberg, L. R. (1984). A failure to replicate the Bem and Allen study of individual differences in cross-situational consistency. *Journal of Personality and Social Psychology, 47,* 1074–1090.

Cheek, J. M. (1982). Aggregation, moderator variables, and the validity of personality tests: A peer-rating study. *Journal of Personality and Social Psychology, 43,* 1254–1269.

Cline, V. B., & Richards, J. M., Jr. (1960). Accuracy of interpersonal perception—A general trait? *Journal of Abnormal and Social Psychology, 60,* 1–7.

Cloyd, L. (1977). Effect of acquaintanceship on accuracy of person perception. *Perceptual and Motor Skills, 44,* 819–826.

Colvin, C. R. (1993a). Childhood antecedents of young-adult judgability. *Journal of Personality, 61,* 611–635.

Colvin, C. R. (1993b). Judgable people: Personality, behavior, and competing explanations. *Journal of Personality and Social Psychology, 64,* 861–873.

Colvin, C. R., Block, J., & Funder, D. C. (1995). Overly positive self-evaluations and personality: Negative implications for mental health. *Journal of Personality and Social Psychology, 68,* 1152–1162.

Colvin, C. R., & Funder, D. C. (1991). Predicting personality and behavior: A boundary on the acquaintanceship effect. *Journal of Personality and Social Psychology, 60,* 884–894.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally.

Cronbach, L. J. (1955). Processes affecting scores on "understanding of others" and "assumed similarity." *Psychological Bulletin, 52,* 281–302.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

Crow, W. J., & Hammond, K. R. (1957). The generality of accuracy and response in interpersonal perception. *Journal of Abnormal and Social Psychology, 54,* 384–390.

Csikszentmihalyi, M., & Larson, R. (1992). Validity and reliability of the Experience Sampling Method. In M. V. de Vries (Ed.), *The experience of psychopathology: Investigating mental disorders in their natural settings* (pp. 43–57). Cambridge, England: Cambridge University Press.

Donahue, E. M., Robins, R. W., Roberts, B. W., & John, O. P. (1993). The divided self: Concurrent and longitudinal effects of psychological adjustment and social roles on self-concept differentiation. *Journal of Personality and Social Psychology, 64,* 834–846.

Dunning, D., Perie, M., & Story, A. L. (1991). Self-serving prototypes of social categories. *Journal of Personality and Social Psychology, 61,* 957–968.

Estes, S. G. (1938). Judging personality from expressive behavior. *Journal of Abnormal and Social Psychology, 33,* 217–236.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Fletcher, G. (1993, October). *Empathic accuracy in married couples.* Paper presented at the meeting of the Society for Experimental Social Psychology, Santa Barbara, CA.

Flink, C., & Park, B. (1991). Increasing consensus in trait judgments through outcome dependency. *Journal of Experimental Social Psychology, 27,* 453–467.

Fuhrman, R. W., & Funder, D. C. (in press). Convergence between self

and peers in the response-time processing of trait-relevant information. *Journal of Personality and Social Psychology.*

Funder, D. C. (1980a). On seeing ourselves as others see us: Self–other agreement and discrepancy in personality ratings. *Journal of Personality, 48,* 473–493.

Funder, D. C. (1980b). The "trait" of ascribing traits: Individual differences in the tendency to trait ascription. *Journal of Research in Personality, 14,* 376–385.

Funder, D. C. (1982). On the accuracy of dispositional vs. situational attributions. *Social Cognition, 1,* 205–222.

Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. *Psychological Bulletin, 101,* 75–90.

Funder, D. C. (1989). Accuracy in personality judgment and the dancing bear. In D. Buss & N. Cantor (Eds.), *Personality research for the 1990's* (pp. 210–223). New York: Springer-Verlag.

Funder, D. C. (1991). Global traits: A neo-Allportian approach to personality. *Psychological Science, 2,* 31–39.

Funder, D. C. (1993a). Judgments as data for personality and developmental psychology: Error versus accuracy. In D. C. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds.), *Studying lives through time: Approaches to personality and development* (pp. 121–146). Washington, DC: American Psychological Association.

Funder, D. C. (1993b). Judgments of personality and personality itself. In K. M. Craik, R. T. Hogan, & R. N. Wolfe (Eds.), *Fifty years of personality psychology* (pp. 207–214). New York: Plenum Press.

Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology, 55,* 149–158.

Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology, 60,* 773–794.

Funder, D. C., & Colvin, C. R. (in press). Congruence of self and others' judgments of personality. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology.* San Diego, CA: Academic Press.

Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52,* 409–418.

Funder, D. C., Kolar, D. C., & Blackman, M. C. (1995). Agreement among judges of personality: Interpersonal relations, similarity, and acquaintanceship. *Journal of Personality and Social Psychology, 69,* 656–672.

Funder, D. C., & Sneed, C. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology, 64,* 479–490.

Funder, D. C., & West, S. G. (1993a). Consensus, self–other agreement, and accuracy in personality judgment: An introduction. *Journal of Personality, 61,* 457–467.

Funder, D. C., & West, S. G. (1993b). Viewpoints on personality: Consensus, self–other agreement, and accuracy in personality judgment [Special issue]. *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.

Gangestad, S. W., Simpson, J. A., DiGeronimo, K., & Beik, M. (1992). Differential accuracy across traits: Examination of a functional hypothesis. *Journal of Personality and Social Psychology, 62,* 688–698.

Gibson, J. J. (1979). *The ecological approach to visual perception.* New York: Harper & Row.

Gifford, R. (1994). A lens-mapping framework for understanding the encoding and decoding of interpersonal dispositions in nonverbal behavior. *Journal of Personality and Social Psychology, 66,* 398–412.

Gifford, R., Fan Ng, C., & Wilkinson, M. (1985). Nonverbal cues in

the employment interview: Links between applicant qualities and interviewer judgments. *Journal of Applied Psychology, 70,* 729–736.

Gifford, R., & Hine, D. W. (1994). The role of verbal behavior in the encoding and decoding of interpersonal dispositions. *Journal of Research in Personality, 28,* 115–132.

Gigerenzer, G. (1991a). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review, 98,* 254–267.

Gigerenzer, G. (1991b). How to make cognitive illusions disappear: Beyond "heuristics and biases." In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 2, pp. 83–115). London: Wiley.

Gigerenzer, G. (1991c). Probabalistic mental models: A Brunswikian theory of confidence. *Psychological Review, 98,* 506–528.

Glanzer, M., & Cunitz, A. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5,* 351–360.

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

Hall, J. A. (1990). *Nonverbal sex differences: Accuracy of communication and expressive style.* Baltimore: Johns Hopkins University Press.

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, England: Cambridge University Press.

Hastorf, A. H., & Bender, I. E. (1952). A caution respecting the measurement of empathic ability. *Journal of Abnormal and Social Psychology, 47,* 574–576.

Havenstein, N. M., & Alexander, R. A. (1991). Rating ability in performance judgments: The joint influence of implicit theories and intelligence. *Organizational Behavior and Human Decision Processes, 50,* 300–323.

Hofstee, W. K. B. (1994). Who should own the definition of personality? *European Journal of Personality, 8,* 149–162.

Hollander, E. P. (1957). The reliability of peer nominations under various conditions of administration. *Journal of Applied Social Psychology, 41,* 85–90.

Hollander, E. P. (1965). Validity of peer nominations in predicting a distant performance criterion. *Journal of Applied Social Psychology, 49,* 434–438.

Ickes, W. (1993). Empathic accuracy. *Journal of Personality, 61,* 587–610.

Jackson, D. N., Neill, J. A., & Bevan, A. R. (1969). Interpersonal judgmental accuracy and bias as a function of degree of acquaintance. *Proceedings of the 77th Annual Convention of the American Psychological Association, 4,* 135–136.

James, W. (1915). *The will to believe and other essays in popular philosophy.* London: Longmans, Green. (Original work published 1897)

John, O. P., & Robins, R. W. (1994a). Accuracy and bias in self-perception: Individual differences in self-enhancement and narcissism. *Journal of Personality and Social Psychology, 66,* 206–219.

John, O. P., & Robins, R. W. (1994b). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality, 61,* 521–551.

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., & Nisbett, R. E. (1971). *The actor and the observer: Divergent perceptions of the causes of behavior.* Morristown, NJ: General Learning Press.

Jourard, S. M. (1971). *Self-disclosure: An experimental analysis of the transparent self.* New York: Wiley.

Jussim, L. (1993). Accuracy in interpersonal expectations: A reflec-

tion–construction analysis of current and classic research. *Journal of Personality, 61,* 637–668.

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

Kane, J. S., & Lawler, E. E., III (1978). Methods of peer assessment. *Psychological Bulletin, 85,* 555–586.

Kenny, D. A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review, 92,* 155–163.

Kenny, D. A. (1993). A coming-of-age for research on interpersonal perception. *Journal of Personality, 61,* 789–807.

Kenny, D. A. (1994). *Interpersonal perception.* New York: Guilford Press.

Kenrick, D. T., & Stringfield, D. O. (1980). Personality traits and the eye of the beholder: Crossing some traditional philosophical boundaries in the search for consistency in all of the people. *Psychological Review, 87,* 88–104.

Kirouac, G., & Dore, F. Y. (1985). Accuracy of the judgment of emotions as a function of sex and level of education. *Journal of Nonverbal Behavior, 9,* 3–7.

Koestner, R., Bernieri, F., & Zuckerman, M. (1989). Trait-specific versus person-specific moderators of cross-situational consistency. *Journal of Personality, 57,* 1–16.

Koestner, R., Bernieri, F., & Zuckerman, M. (1994). Self–peer agreement as a function of two kinds of trait relevance. *Social Behavior and Personality, 22,* 17–30.

Kolar, D. W., Funder, D. C., & Colvin, C. R. (in press). Comparing the accuracy of personality judgments by the self and knowledgable others. *Journal of Personality.*

Koppenaal, L., & Glanzer, M. (1990). An examination of the continuous distractor task and the "long-term recency effect." *Memory & Cognition, 18,* 183–195.

Kruglanski, A. W. (1989). The psychology of being "right": The problem of accuracy in social perception and cognition. *Psychological Bulletin, 106,* 395–409.

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

Kuiken, D. (1981). Non-immediate language style and inconsistency between private and expressed evaluations. *Journal of Experimental Social Psychology, 17,* 183–196.

Lanning, K. (1988). Individual differences in scalability: An alternative conception of consistency for personality theory and measurement. *Journal of Personality and Social Psychology, 55,* 142–148.

Lester, D. (1991). Accuracy of recognition of genuine versus simulated suicide notes. *Personality and Individual Differences, 12,* 765–766.

Levesque, M. J., & Kenny, D. A. (1993). Accuracy of behavioral predictions at zero acquaintance: A social relations analysis. *Journal of Personality and Social Psychology, 65,* 1178–1187.

Lewin, A. Y., & Zwany, A. (1976). Peer nominations: A model, literature critique and a paradigm for research. *Personnel Psychology, 29,* 423–447.

Lopes, L. L. (1991). The rhetoric of irrationality. *Theory and Psychology, 1,* 65–82.

MacLellan, B. N., & Jackson, D. N. (1985). Accuracy and consistency in the development of social perception. *Developmental Psychology, 21,* 30–36.

Marangoni, C., Garcia, S., Ickes, W., & Teng, G. (1995). Empathic accuracy in a clinically relevant setting. *Journal of Personality and Social Psychology, 68,* 854–869.

Markus, H., Smith, J., & Moreland, R. L. (1985). Role of the self-concept in the perception of others. *Journal of Personality and Social Psychology, 49,* 1494–1512.

Mayfield, E. C. (1972). Value of peer nominations in predicting life insurance sales performance. *Journal of Applied Psychology, 56,* 319–323.

McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review, 90,* 215–238.

McCrae, R. R. (1994). The counterpoint of personality assessment: Self-reports and observer ratings. *Assessment, 1,* 159–172.

Moskowitz, D. S., & Schwarz, J. C. (1982). Validity comparison of behavior counts and ratings by knowledgeable informants. *Journal of Personality and Social Psychology, 42,* 518–528.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* New York: Prentice Hall.

Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality structure. *Journal of Personality and Social Psychology, 4,* 681–691.

Ozer, D. J. (1989). Construct validity in personality assessment. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 224–234). New York: Springer-Verlag.

Park, B., & Judd, C. M. (1989). Agreement on initial impressions: Differences due to perceivers, trait dimensions, and target behaviors. *Journal of Personality and Social Psychology, 56,* 493–505.

Paulhus, D. L., & Bruce, M. N. (1992). The effect of acquaintanceship on the validity of personality impressions: A longitudinal study. *Journal of Personality and Social Psychology, 63,* 816–824.

Paunonen, S. V. (1989). Consensus in personality judgments: Moderating effects of target–rater acquaintanceship and behavior observability. *Journal of Personality and Social Psychology, 56,* 823–833.

Paunonen, S. V., & Jackson, D. N. (1987). Accuracy of interviewers and students in identifying the personality characteristics of personnel managers and computer programmers. *Journal of Vocational Behavior, 31,* 26–36.

Reise, S. P., & Waller, N. G. (1993). Traitedness and the assessment of response pattern scalability. *Journal of Personality and Social Psychology, 65,* 143–151.

Reisman, J. M. (1990). Intimacy in same-sex friendships. *Sex Roles, 23,* 65–82.

Robins, R. W., & John, O. P. (in press). The quest for self-insight: Theory and research on the accuracy of self-perception. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology.* San Diego, CA: Academic Press.

Rorer, L. G. (1991). Some myths of science in psychology. In D. Cicchetti & W. Grove (Eds.), *Thinking clearly about psychology: Essays in honor of Paul E. Meehl* (pp. 61–87). Minneapolis: University of Minnesota Press.

Ross, L. (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 174–214). New York: Academic Press.

Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology.* New York: McGraw-Hill.

Rothbart, M., & Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology, 50,* 131–142.

Scherer, K. R. (1978). Personality inference from voice quality: The loud voice of extraversion. *European Journal of Social Psychology, 8,* 467–487.

Schmidt, F. L., Hunter, J. E., Croll, P. R., & McKenzie, R. C. (1983). Estimation of employment test validity by expert judgment. *Journal of Applied Psychology, 68,* 590–601.

Schneider, D. J., Hastorf, A. H., & Ellsworth, P. C. (1979). *Person perception* (2nd ed.). Reading, MA: Addison-Wesley.

Sedikides, C. (1993). Assessment, enhancement, and verification determinants of the self-evaluation process. *Journal of Personality and Social Psychology, 65,* 317–338.

Sedikides, C., & Skowronski, J. J. (1993). The self in impression for-

mation: Trait centrality and social perception. *Journal of Experimental Social Psychology, 29,* 347–357.

Sherrod, D. R. (1989). The influence of gender on same-sex friendships. In C. Hendrick (Ed.) *Close relationships: Review of personality and social psychology* (Vol. 10). Newbury Park, CA: Sage.

Sillars, A. L., & Scott, M. D. (1983). Interpersonal perception between intimates: An integrative review. *Human Communication Research, 10,* 153–176.

Simpson, J., Ickes, W., & Blackstone, J. (1994). *When the head protects the heart: Motivated inaccuracy in dating relationships.* Manuscript submitted for publication.

Snyder, M. (1987). *Public appearances, private realities: The psychology of self-monitoring.* New York: Freeman.

Snyder, M., & Ickes, W. (1985). Personality and social behavior. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 2, pp. 883–948). New York: Random House.

Spain, J. (1994). *Personality and daily life experience: Evaluating the accuracy of personality judgments.* Unpublished doctoral dissertation, University of California, Riverside.

Stiff, J. B., Kim, H. T., & Ramesh, C. N. (1992). Truth biases and aroused suspicion in relational deception. *Communication Research, 19,* 326–345.

Strasburger, E. L., & Jackson, D. N. (1977). Improving accuracy in a clinical judgment task. *Journal of Consulting and Clinical Psychology, 45,* 303–309.

Swann, W. B., Jr. (1984). Quest for accuracy in person perception: A matter of pragmatics. *Psychological Review, 91,* 457–477.

Taft, R. (1955). The ability to judge people. *Psychological Bulletin, 52,* 1–23.

Taft, R. (1966). Accuracy of empathic judgments of acquaintances and strangers. *Journal of Personality and Social Psychology, 3,* 600–604.

Tagiuri, R., & Petrullo, L. (1958). *Person perception and interpersonal behavior.* Stanford, CA: Stanford University Press.

Tellegen, A. (1988). The analysis of consistency in personality assessment. *Journal of Personality, 56,* 621–663.

Waters, L. K., & Waters, C. W. (1970). Peer nominations as predictors of short-term sales performance. *Journal of Applied Psychology, 54,* 42–44.

Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology, 57,* 120–128.

Westbrook, M. (1974). Judgment of emotion: Attention vs. accuracy. *British Journal of Social and Clinical Psychology, 13,* 383–389.

Zuckerman, M., Bernieri, F., Koestner, R., & Rosenthal, R. (1989). To predict some of the people some of the time: In search of moderators. *Journal of Personality and Social Psychology, 57,* 279–293.