Comparing the Accuracy of Personality Judgments by the Self and Knowledgeable Others

David W. Kolar and David C. Funder
University of California, Riverside
C. Randall Colvin
Northeastern University

ABSTRACT In this article we compare the accuracy of personality judgments by the self and by knowledgeable others. Self- and acquaintance judgments of general personality attributes were used to predict general, videotaped behavioral criteria. Results slightly favored the predictive validity of personality judgments made by single acquaintances over self-judgments, and significantly favored the aggregated personality judgments of two acquaintances over self-judgments. These findings imply that the most valid source for personality judgments that are relevant to patterns of overt behavior may not be self-reports but the consensus of the judgment of the community of one's peers.

Personality judgments about the self and about others are an important part of our daily lives. They help to determine our behavior and how we think and feel about ourselves and others. Researchers who study personality judgments have focused both on the process and the content of these judgments. Research on the process of personality judgment tends to focus on potential errors (Funder, 1987), while more recent research on the content of personality judgments, the focus of the present

This research was supported by National Institute of Mental Health Grant RO1-MH42427 to David Funder. We are grateful for the comments of Dianne Nilsen and Daniel Ozer, and members of the Riverside Accuracy Project (Melinda Blackman, Carl Sneed, Jana Spain, and Leslie Wiehl). Address correspondence to David C. Funder, Department of Psychology, University of California, Riverside, CA 92521-0426. Internet: Funder@ucrac1.ucr.edu.

Journal of Personality 64:2, June 1996. Copyright © 1996 by Duke University Press.

article, evaluates judgmental accuracy (Colvin & Funder, 1991; Funder, 1995; Funder & West, 1993).

Much of the evidence for the accuracy of personality judgments about the self and others comes from research on interjudge agreement. This research has established convincingly that agreement is generally quite good (Kenrick & Funder, 1988). Researchers have also found that more visible traits tend to be judged with better self-other and interjudge agreement (Funder & Dobroth, 1987; Paunonen, 1989; Zuckerman, Bernieri, Koestner, & Rosenthal, 1989), and that close acquaintances provide judgments that agree better with each other and with self-judgments than do the judgments of relative strangers (Colvin & Funder, 1991; Funder & Colvin, 1988; Paunonen, 1989). Using interjudge agreement as a criterion for accuracy, this literature indicates that personality judgments, both from the self and from knowledgeable acquaintances, tend to be accurate.

A natural next step for accuracy research is to determine if one particular type of judge tends to be more accurate than another. One obvious place to begin is with the question "Whose perspective on personality is more accurate, the view held by the self or the view held by knowledgeable others?" So far, research results on this question have tended to be less than conclusive. The purpose of the present article is to present some new evidence relevant to this issue.

First, we will examine some theoretical considerations and empirical evidence indicating that self-judgments of personality might be more accurate than judgments by others. Second, we will examine some theoretical considerations and empirical evidence that indicate personality judgments made by close acquaintances might be more accurate than self-judgments. Third, we will discuss research that has begun to address self- versus other accuracy in personality judgments. Fourth, we will examine and evaluate the results of our own, new study on this topic. Finally, we will discuss the implications of our results for research on personality judgment.

Behavior as a Criterion for Accuracy

We have mentioned that by far the most common criterion for accuracy in personality judgments has been interjudge agreement. Sometimes self-reports have been used as criteria for the accuracy of peers' reports (Paunonen, 1989), sometimes peers' reports have been used as criteria for the accuracy of self-reports (Hase & Goldberg, 1967; Watson &

Clark, 1991), and sometimes peer and self-reports have been used as criteria for the accuracy of each other (Funder & Colvin, 1988; Funder & Dobroth, 1987).

However, agreement is not satisfactory as an ultimate criterion (Colvin & Funder, 1991; Funder, 1987; Kenny, 1991). For one thing, agreement does not necessarily imply accuracy (although accuracy does imply agreement); two judgments can agree with each other perfectly, yet both be perfectly wrong. A more specific shortcoming becomes obvious in the context of the research question raised by the present study: The traditional criterion of self-other agreement will clearly not serve if we wish to examine the *relative accuracy* of self- and other judgments of personality. This makes it necessary for us to discuss and to use an alternative criterion for judgmental accuracy that is becoming increasingly popular: behavioral prediction (e.g., Colvin & Funder, 1991; John & Robins, 1994; Nilsen, 1991). A judgment of personality is deemed accurate to the extent it can predict independent evaluations of the behavior of the person judged (Funder & West, 1993). This approach is congruent with a traditional validation strategy of personality psychology that evaluates personality measures by assessing the degree to which they are predictive of broad classes of performance or more specific behavioral criteria (Wiggins, 1973).

Why Self-Judgments May Be More Accurate

Our intuitions as well as widespread viewpoints within personality and social psychology seem to indicate that the self is in a uniquely advantaged and perhaps even privileged position to judge its own personality accurately. We know how we act across different settings, we have a unique perspective on our private experiences, and in general we have access to more information about ourselves than any observer could possibly have. Indeed, as Nilsen (1991) points out, "[T]he idea that people don't know how they behave seems quite outlandish" (p. 10).

Consistent with this view, when personality psychologists want to know what someone is like, they usually simply ask that person. Self-report questionnaires are by far the favorite source of data within personality psychology (Hofstee, 1994; Wiggins, 1973). Social psychologists manifest a parallel tendency to favor the self's point of view. For instance, when they examine the classic "actor-observer" effect, social psychologists tend to ascribe greater credibility to the actor's typical self-report that his or her behavior was determined by situational factors

than to the observer's typical attribution of the behavior to dispositional properties of the actor (Jones, 1979; Monson & Snyder, 1977).

Research on the accuracy of self-judgments indicates not only that self-judgments often agree well with judgments made by others, but also that they can be efficient at predicting behavior. Researchers have shown that self-judgments can predict such outcomes as insomnia (Coates et al., 1982), smoking (Morrell, King, & Martin, 1986), and alcohol consumption (Sobell & Sobell, 1975), among others.

Why Judgments Made by Others May Be More Accurate

Other theoretical considerations raise the possibility that in some respects individuals may actually be in a poorer position to evaluate their own behavior and personality than are close friends and acquaintances. The superiority of self-report can be doubted for several reasons.

The first is that individuals may present themselves in an unrealistically positive light when describing themselves on personality measures, either because they are trying to deceive the personality tester or because they are deceiving themselves (see Greenwald, 1980; Hofstee, 1994; Holzbach, 1978; John & Robins, 1993; Kenny, 1994). This possibility has received a fair amount of attention, and a number of strategies have been developed for minimizing or correcting the effect of "response bias" (Paulhus, 1991). More recently, John and Robins (1994) have presented evidence showing that some individuals (specifically, narcissists) are particularly likely to self-enhance. Clinical psychologists also have long believed that individuals are frequently unwilling to report or are even unaware of their own shortcomings.

A second reason to doubt the perspicacity of self-report has received less attention. That reason is perspective. Individuals may be in a poor position to see their own consistent personality attributes, both because of the literal angle from which they view themselves and the circumstances under which they do the viewing. This can be seen by again looking at the actor-observer literature. Although this literature traditionally has emphasized the apparent informational advantages of the actor, other aspects of the classical analysis can be taken to derive advantages that might accrue to an observer.

Specifically, it has been pointed out by Jones and Nisbett (1971) and others (e.g., Storms, 1973) that the observational perspective of the actor is on the surrounding environment, whereas from the perspective

of an observer, the actor is located in a perceptual field alongside other actors with whom the individual is then naturally compared (Heider, 1958). Experiments that have manipulated this aspect of observational perspective, while holding constant other aspects (e.g., information about the actor's behavior in other situations), find that taking the observer's perspective is sufficient to cause one to make attributions for behavior that are more dispositional, and less situational, than taking the actor's perspective. Actors experimentally induced to take an observer's perspective on their own behavior have been shown to make more dispositional attributions when the perspective was manipulated by video cameras (Storms, 1973), mirrors (Duval & Wicklund, 1973), explicit instructions to take the observers' perspective (Frank & Gilovich, 1989), and the sheer passage of time (Moore, Sherrod, Liu, & Underwood, 1979; Peterson, 1980). As Frank and Gilovich (1989) summarize the literature, "All of these manipulations serve to make actors more aware of themselves and their actions and thus also lead them to attribute their behavior more dispositionally" (p. 402). The implication for the present context is that the same factors that make dispositional influences on behavior more salient from the perspective of an observer than from the perspective of an actor might also cause the relevant dispositions to be more accurately judged from that perspective.

Of course, the literature just summarized focuses primarily on comparing attributions about specific behaviors made by the self and others, while the focus of the present research is on comparing judgments of personality made by the self and others. Because individuals making judgments of personality tend to describe themselves or others as they generally are (i.e., over long periods of time), one might suspect that both sets of judgments (self- and other) would tend to be dispositional and therefore the differences in perspective would become less marked than the attribution literature suggests.

In addition to these theoretical considerations, empirical evidence also suggests that the observer's perspective might be advantageous for making accurate judgments of individual differences in personality and behavior. The literature on "objective self-awareness" has demonstrated that judgments of personality made from an external vantage point are not only more dispositional, but likely to have greater predictive validity as well (Scheier, Buss, & Buss, 1976).

Indeed, Hofstee (1994), on the basis of theoretical and empirical considerations such as summarized above, recommends that when self-report *is* used, "the writing of personality questionnaires [should be

done] in the third person singular" (p. 159). This practice is intended to improve the accuracy of self-judgment by forcing one to take the psychological position of an outside observer on oneself. But an even better practice, Hofstee claims, "is to turn to third persons of flesh and blood in assessing personality" (p. 159).

Comparing the Ability of Different Sources to Predict Behavior

Some of the research cited previously uses behavioral prediction in an attempt to evaluate the accuracy of personality judgments made by the self and others. In this section, we will briefly review some research that has used behavioral prediction to *compare* the accuracy of different sources of judgment.

In a review of the research literature, Shrauger and Osberg (1981) evaluated several studies that compared the ability of self-reports with other assessment techniques (e.g., interviews, historical data, psychological test data) to predict specific criterion behavior (e.g., GPA, course grades, college major). Their conclusions as to what source did better were mixed and depended on the criterion. They concluded that self-reports were superior to other assessment data when attempting to predict vocational choice or peer ratings of the target. Self-reports were slightly favored when attempting to predict psychotherapy outcomes and academic achievement. Finally, neither self-report nor other assessment data were deemed superior when attempting to predict post-hospital adjustment and job performance.

However, most of the studies that these authors evaluated compared self-judgments with some kind of test data (e.g., TAT score, SAT score, inventory score) and not with judgments of personality made by close acquaintances. Those studies that did directly compare the accuracy of self-judgments with judgments by other people were split in their results. Some favored self-judgments, while others favored judgments made by others.

More recent research that has directly compared the accuracy of self- and other judgments includes work by John and Robins (1994). These authors found that the other participants in a group discussion ranked each actor's contribution to the group more accurately (compared with highly reliable criterion rankings by psychologists) than did the actors themselves. Similar results have been reported by Levesque and Kenny (1993). Another recent study by Nilsen (1991) found that observers'

ratings were more valid than self-ratings for predicting managerial performance.

Using Personality Judgments to Predict Psychologically Meaningful Behavior

Each of the studies cited in the preceding section makes a useful contribution. Still, as Kenny (1994) notes, in the aggregate they comprise "not much evidence concerning the relative validity of peer perception versus other perception" (p. 194). Most of these studies are of limited implication in that they gather self- and other judgments of performance rather than of more general attributes of personality. On the criterion side, most of these studies focus on specific behaviors such as academic achievement, alcohol consumption, insomnia, managerial performance, or contribution to a group.

While these predictors and criteria are no doubt important to understand, the next step for researchers should be to examine the relative ability of the self and of peers to predict behaviors at a more psychologically implicative level. That is, beyond specific performance and other relatively narrow criteria, we might want to be able to predict the more general attributes of an individual's personality and behavior. Does the person have a high aspiration level, behave in a sympathetic manner, or overreact to minor frustrations? These kinds of general patterns of behavior are often important to understand in ourselves and in other people and are quite relevant to personality. Behaviors like these are the focus of the present research.

In the next section, we present a new study on the accuracy of personality judgments by the self and others. We will compare the ability of judgments of personality to predict participants' psychologically meaningful behaviors, as coded from several videotaped interpersonal interactions. We will in this way compare the predictive validity of self-reports with judgments by a single close acquaintance and judgments yielded by the aggregate of two acquaintances.

METHOD

Participants and Prior Studies

The present analyses are based on a large and complex data set that has yielded several prior studies. One hundred forty undergraduates at Harvard University, 70 of each sex, were videotaped in three laboratory situations and also

provided self-descriptions of personality through a variety of formats. The students' videotaped behavior was coded as described below. The personalities of these students were described by friends and roommates. The exact *N* varies across specific analyses due to occasionally incomplete data.

The present study focuses upon relationships between self- and peer descriptions of personality and the behavior as coded from the videotapes. Prior studies have examined the relationship between self- and peer descriptions and measures of participants' "social acuity" (Funder & Harris, 1986), individual differences in "judgability" (Colvin, 1993), differences between traits that yield better and poorer interjudge agreement (Funder & Dobroth, 1987), the effect of level of acquaintanceship on interjudge agreement (Funder & Colvin, 1988) and on the ability to predict measures of personality and behavior (Colvin & Funder, 1991), the consistency of behavior across the three laboratory situations as well as between the laboratory and daily life (Funder & Colvin, 1991), how lay judges use specific behavioral cues in personality judgment (Funder & Sneed, 1993), and how people who self-enhance their personality judgments behave in interpersonal settings (Colvin, Block, & Funder, in press).

Personality Judgments

Personality judgments were obtained using the California Q-sort (Block, 1961/1978), as slightly modified by Bem and Funder (1978) for use with nonprofessionals. A judge of personality arranges the 100 items of the Q-set (e.g., "is critical, skeptical, not easily impressed") into a forced, 9-step distribution ranging from "highly uncharacteristic" (1) to "highly characteristic" (9) of the individual being described. Each participant completed a self Q-sort. Also, two close acquaintances of nearly every participant completed a Q-sort of the participant. It is important to note that these close acquaintances completed the Q-sorts using only information from past experiences with the participants. These acquaintances *did not* view the videotapes of their friends that were the source of the behavioral criterion. Further details of the Q-sorting procedure were reported by Funder and Colvin (1988).

Behavioral Coding

The behavior of each participant was recorded on videotape in three situations, denoted Time 1, Time 2, and Debate. At Time 1, the participant was introduced to a person of the opposite sex whom he or she had never met before but who was also participating in the study for the first time, and then was left alone to converse with that person for 5 minutes. The Time 2 session, which occurred a few weeks later, was the same as at Time 1 except that both participants, who again had never met before, were attending their second experimental

session. The Debate occurred a few minutes after Time 2, with the same partner. Participants were assigned sides in a debate on capital punishment and their debate was taped for 5 minutes.

In order to analyze the behavior on these videotapes, we developed a behavioral coding scheme. The coding of behavioral data is a cumbersome task that can be done in a nearly infinite number of ways. Options for coding such data range from specific counts of small acts (eyebrow twitches, elbow lifts) to highly impressionistic judgments of the meaning of behavior (Cairns & Green, 1979; Fiske, 1979). There is no one correct coding scheme for all types of behavioral data. Rather, the coding scheme ultimately used should be determined by one's research goals (Bakeman & Gottman, 1986).

In the present research, our central goal was to extract units of behavior that would be psychologically meaningful in their own right and that would be relatively likely to be relevant to our participants' general personality and behavior at large. The type of behavior in which we are interested is similar to what Tolman (1932) called "purposive behavior." Tolman believed that "behavior-acts, though no doubt in complete one-to-one correspondence with the underlying molecular facts of physics and physiology, have, as 'molar' wholes, certain emergent properties of their own. And it is these, the molar properties of behavior-acts, which are of prime interest to us as psychologists" (p. 7).

More recently, Mischel (1973) provided a clear exposition of the advantages of assessing behaviors at the level of analysis at which they have intrinsic meaning:

> Early versions of behaviorism [employed] . . . simplistic definitions [of behavior] in terms of clearly delineated motor "acts" (such as bar press). . . . More recent versions of behavior theory, moving from cat, rat, and pigeon confined in the experimenter's apparatus to people in exceedingly complex social situations, have extended the domain of studied behavior much beyond motor acts and muscle twitches; they seek to encompass what people do cognitively, emotionally, and interpersonally, not merely their arm, leg, and mouth movements. Now the term "behavior" has been expanded to include virtually anything that an organism does, overtly or covertly, in relation to extremely complex social and interpersonal events. . . . Such categories go considerably beyond self-evident behavior descriptions . . . [and involve] inferences about the subject's intention and abstractions about behavior, rather than mere physical description of actions and utterances. (p. 268)

In a similar spirit, the research project of which the present study is part developed a behavioral Q-set (BQ) to assess behavior at the level of analysis at which it is not only intrinsically meaningful, as described by Mischel, but also is most likely to show personality-relevant consistency (Cairns & Green, 1979;

Funder & Colvin, 1991; for a similarly motivated use of a different behavioral Q-set, see Langston & Cantor, 1989). Its 62 items include general behavioral descriptors such as "speaks in a loud voice," "behaves in a fearful or timid manner," and "has high enthusiasm and high energy level." An average of six coders viewed each videotape and described the behavior of a participant by arranging the 62 items into a forced, 9-step distribution ranging from "not at all characteristic" (1) to "highly characteristic" (9) of the participant in the interaction. Each coder coded the behavior of each participant only once, to maintain the independence of data between participants and across sessions. The median composite reliability across the three sessions for the 62 items was .64. A principal components analysis followed by an orthogonal rotation of the 62 Behavioral Q-sort items yielded four latent factors, labeled nervous withdrawal, domineeringness, serious intelligence, and heterosexuality.[1] This analysis was reported in detail by Colvin and Funder (1991). Further details of the videotaping procedure and the development, implementation, and reliability of the Behavioral Q-sort are reported in Funder and Colvin (1991).

In general, the BQ items showed good cross-situational consistency (Funder & Colvin, 1991). For the purposes of the present study, each of the behavioral items as well as the four behavioral factor scores was aggregated across the three videotaped sessions. The behavioral criterion was aggregated in this way in order to provide an appropriately general criterion for the validity of the general personality judgments (Fishbein & Ajzen, 1974).

Cognate Items

Forty-one of the BQ items were written to almost exactly parallel "cognate" items in the personality Q-sort. For example, the CQ (personality) item "is cheerful" has a cognate BQ (behavioral) item that reads "behaves in a cheerful manner."[2] These 41 items are used for most of the analyses presented in the Results section.

RESULTS

We compared the predictive validity of personality judgments by the self with those by acquaintances using two methods, one nomothetic and the other idiographic. The first analysis compares the ability of per-

1. We also conducted a principal factor analysis followed by an oblique rotation. The factor structure and loadings were essentially the same as the principal components analysis with an orthogonal rotation.
2. Note that the CQ item, "has a wide range of interests," has two BQ cognates.

sonality Q-sorts from the two sources to predict each of the 41 specific behavioral items that have California Q-set (personality) cognates, as well as the four nomothetic behavioral factors derived from the behavioral Q-sort. The second analysis compares the ability of individual judges of each type to provide personality Q-sorts that correlate highly with each target's individual behavioral Q-sort across the 41 items that have equivalents in the two Q-sets (Funder & Colvin, 1991). To describe this important distinction another way, the first analysis focuses on the predictability of individual differences between targets on each of the 41 nomothetic behavioral items and on the four nomothetic behavioral factors. The second analysis focuses on the predictability of the idiographic, differential salience of each of the 41 behavioral items, compared to each other, within each target (Block, 1961/1978).

Nomothetic Analyses

For the nomothetic analyses we simply computed the correlation between each of the 41 matched personality Q-sort items and its behavioral cognate, or equivalent item. These correlations were calculated separately using the personality descriptions provided by the self, by each acquaintance, and by the aggregate (average) of the two acquaintances. The results appear in Table 1.

In this analysis, it can be seen that in general the ratings of a single acquaintance have slightly more predictive validity than a self-rating, with the average correlation for the (arbitrarily designated) first acquaintance being significantly greater than the average for the self, and that for the second acquaintance being nonsignificantly greater.[3] The predictive validity of the aggregate of the two acquaintances' ratings, which benefits from the enhanced reliability that accrues to aggregated ratings, was on average significantly greater than that of the self-rating and of either acquaintance individually.

A second, related nomothetic analysis, reported in Table 2, reaches the same conclusion. In this analysis, scores for each of the commonly used Big Five personality factors were derived from the personality Q-sort, using the same 41 CQ-BQ cognates employed in the preced-

3. A paired-comparison t test was used to determine if the average correlations were significantly different in Table 1; all averaging of correlations employed Fisher's r-to-z conversion.

Table 1
Correlation between the 41 Corresponding California Q-Sort (CQ) and Behavioral Q-Sort (BQ) Items

CQ item/BQ item	Self ($n = 139$)	Peer 1 ($n = 137$)	Peer 2 ($n = 127$)	C Peer ($n = 137$)
Has a wide range of interests/Seems interested in what partner says	−.07	.13	.09	.16
Behaves in an assertive fashion/Tries to control interaction	.13	.26*	.21*	.24*
Genuinely submissive; accepts domination comfortably/Dominates interaction (−)	.06	.27*	.24*	.18*
Calm, relaxed in manner/Appears to be relaxed and comfortable	.17*	.01	−.04	−.01
Has social poise and presence/Exhibits social skills	.19*	.29*	.23*	.31*
Is emotionally bland/Is reserved and unexpressive	.26*	.23*	.27*	.32*
Has a wide range of interests/Discusses unusually large number of topics	.08	.06	.10	.06
Is critical, skeptical, not easily impressed/Expresses skepticism or cynicism	.29*	.29*	.06	.24*
Is a talkative individual/Is talkative	.10	.24*	.28*	.30*
Is basically anxious/Expresses insecurity or sensitivity	.01	.04	−.10	−.04
Anxiety and tension find outlet in bodily symptoms/Shows physical signs of anxiety	.00	.01	−.04	−.04

Table 1
Continued

CQ item/BQ item	Self ($n = 139$)	Peer 1 ($n = 137$)	Peer 2 ($n = 127$)	C Peer ($n = 137$)
Appears to have high intellectual capacity/ Exhibits high degree of intelligence	.35*	.19*	.36*	.35*
Behaves in a sympathetic or considerate manner/Expresses sympathy toward partner	.22*	.21*	.12	.21*
Initiates humor/Initiates humor	.09	.23*	.16	.23*
Seeks reassurance from others/Seeks reassurance from partner	.01	.02	.13	.07
Shows condescending behavior to others/Exhibits condescending behavior	.13	.20*	.16	.18*
Tends to arouse liking and acceptance/Partner seems to like him or her	.03	.10	.17	.17*
Is turned to for advice and reassurance/Seeks advice from partner (−)	−.10	.06	−.06	.08
Regards self as physically attractive/Appears to regard self physically attractive	.19*	.44*	.23*	.41*
Overreactive to minor frustrations, irritable/ Acts in an irritable fashion	−.13	.17*	.05	.14
Has warmth; capacity for close relationships; compassionate/ Expresses warmth	.05	.22*	−.08	.12

Table 1
Continued

CQ item/BQ item	Self ($n = 139$)	Peer 1 ($n = 137$)	Peer 2 ($n = 127$)	C Peer ($n = 137$)
Is vulnerable to real or fancied threat; fearful/ Behaves in a fearful manner	.14	.10	.01	.06
Is facially and/or gesturally expressive/ Is expressive in face, voice, or gesture	.26*	.26*	.16	.25*
Engages in personal fantasy and daydreams/ Expresses interest in fantasy or daydream	.03	.19*	.11	.18*
Has a readiness to feel guilt/Expresses guilt	.01	.12	.06	.12
Keeps people at a distance, avoids relationships/Keeps partner at a distance	.16	.17	.22*	.26*
Genuinely values intellectual matters/Shows interest in intellectual matters	.31*	.34*	.40*	.41*
Emphasizes being with others; gregarious/ Seems to enjoy interaction with partner	.15	.20*	.06	.18*
Is an interesting, arresting person/Says or does interesting things	.01	−.06	.06	−.01
Judges self and others in conventional terms/ Shows interest in conventional ways	.08	.02	.00	.01
Has high aspiration level for self/Displays ambition	.23*	.25*	.32*	.33*
Perceives many contexts in sexual terms/Views interaction as sexual encounter	.13	.06	.18*	.13

Table 1
Continued

CQ item/BQ item	Self ($n = 139$)	Peer 1 ($n = 137$)	Peer 2 ($n = 127$)	C Peer ($n = 137$)
Feels cheated and victimized by life/Expresses self-pity	−.05	.23*	.19*	.28*
Interested in members of opposite sex/Interested in partner as member of opposite sex	.09	−.16	.05	−.07
Is cheerful/Behaves in a cheerful manner	.24*	.21*	.31*	.32*
Concerned with philosophical problems/Discusses philosophical issues with interest	.21*	.08	.21*	.15
Is sex-typed (masculine/feminine)/Behaves in masculine/feminine style	.30*	.26*	.28*	.32*
Tends to proffer advice/Offers advice to partner	.09	.02	.05	.05
Is verbally fluent/Speaks fluently and expresses ideas well	.18*	.23*	.16	.25*
Is power oriented/Demonstrates interest in topics related to power	−.02	.02	.12	.08
Has a rapid personal tempo; behaves and acts quickly/Speaks quickly	.23*	.22*	.20*	.28*
Average correlation	$.12_a$	$.16_b$	$.14_{ab}$	$.18_c$

Note. Average correlations not sharing a subscript are significantly different from each other at the .05 level or better. Correlations are not independent. Peer 1 is Informant 1; Peer 2 is Informant 2; C Peer is the composite of the two informants.
*$p < .05$.

Table 2
Predicting Four Behavioral Factors from the Big Five

	Source of personality description			
Behavioral factors	Self ($n = 139$)	Peer 1 ($n = 135$)	Peer 2 ($n = 125$)	Combined peer ($n = 137$)
Nervous withdrawal	$.27_{a}$	$.37*_{ab}$	$.36*_{a}$	$.44*_{b}$
Domineeringness	$.33*_{a}$	$.48*_{bc}$	$.38*_{ab}$	$.47*_{c}$
Serious intelligence	$.37*_{a}$	$.50*_{b}$	$.45*_{ab}$	$.51*_{b}$
Heterosexuality	$.16_{a}$	$.32*_{ab}$	$.26_{ab}$	$.32*_{b}$

Note. Multiple correlations within each row not sharing a subscript are significantly different from each other at $p < .10$, two-tailed, or $p < .05$, one-tailed. Combined peer is the composite of the two informants.
$*p < .05$.

ing analysis (McCrae, Costa, & Busch, 1986).[4] These five personality factors were then combined in a simultaneous multiple regression to predict each of the four behavioral factors. All four of the behavioral factors were better predicted by the aggregated acquaintances' ratings than by the self-ratings. Moreover, as also can be seen in Table 2, all four behavioral factors were also predicted better by *each* of the individual acquaintances' ratings. We believe it is this general pattern that is interesting, but it may be worth noting that several of the differences between the multiple correlations were statistically significant or approached significance as well.[5]

Idiographic Analyses

A second way to look at these data is to compare personality and behavioral Q-sorts idiographically. Using the 41 cognate CQ-BQ items, one

4. Factor scores were computed in the following way. Using factor loadings from McCrae, Costa, and Busch (1986), we determined a threshold factor loading above which an item was to be deemed relevant to a particular factor. The threshold loadings were: Neuroticism, .40; Extraversion, .35; Openness, Agreeableness, and Conscientiousness, all .30. (Varying thresholds were employed across the factors to equalize the number of relevant items for each at approximately 15.) Factor scores were calculated by unit weighting; i.e., item scores were summed across relevant items to derive factor scores.

5. This was computed using a test for the significance of differences between dependent correlations (Roscoe, 1975).

Table 3
Average Profile Correlations across 41 Cognate Items between Personality and Behavioral Descriptions

	Source of personality description			
Individual profiles	Self ($n = 139$)	Peer 1 ($n = 137$)	Peer 2 ($n = 127$)	Combined peer ($n = 137$)
Average correlations	$.27_a$	$.29_a$	$.28_a$	$.33_b$

Note. Average profile correlations not sharing a subscript are significantly different from each other at the .05 level or better. Combined peer is the composite of the two informants.

can compute several "profile" predictive accuracy scores individually for each target.

Each individual accuracy score is simply the correlation coefficient across the 41 pairs of items for each participant, one item in each pair being from the behavioral Q-sort, and the other item being from a personality Q-sort of the participant provided by the self or by one or both acquaintances. The next step is to compute the average idiographic correlation, across the approximately 140 participants, derived from personality Q-sorts from each source. These averages can be compared to each other using a standard paired comparison t test.[6] The results appear in Table 3.

The results of the idiographic analyses echo the nomothetic results. When the acquaintances' predictions are considered one at a time they perform slightly but not significantly better than those of the self. But the two acquaintances' average personality Q-sorts correlated significantly better with the target's behavioral Q-sort than did the targets' own personality Q-sorts. Although computed in a very different manner, this finding from the idiographic analyses is almost exactly parallel to that obtained from the nomothetic analyses reported above.

Possible Reasons for Differences in Accuracy

Previous studies in this research program have analyzed how specific properties of items in the personality and behavioral Q-sorts can affect their correlates (Funder & Colvin, 1988, 1991; Funder & Dobroth,

6. All averaging of idiographic correlations employed the r-to-z conversion.

1987). It is possible that item properties could also affect how accurately those items are judged by the self and others. To test this possibility, we gathered ratings of the favorability and observability of each of the 62 behavioral Q-sort items. These ratings, done by independent panels of eight judges, had aggregate (alpha) reliabilities of .92 and .82, respectively.

Ratings of the "operant/respondent" properties of each item were obtained previously by Funder and Colvin (1991). "Operant" behaviors are behaviors that are characteristically "emitted" by people, such as "behaves in a cheerful manner." "Respondent" behaviors are behaviors that are "elicited" by specific situational stimuli, such as "expresses sympathy to partner." See Colvin and Funder (1991) for a more detailed discussion of this distinction.

To determine if these properties of the behavioral items could explain the differences in predictive efficacy between the self- and acquaintance judgments, we computed a difference score by subtracting acquaintance predictions from the self-predictions across the items in Table 1.[7] We then correlated this difference score with the ratings for the different item properties.

Across the 41 cognate CQ-BQ items, none of the behavioral item properties we assessed correlated significantly with any of the differences in predictive efficacy between the self- and acquaintance judgments that are the focus of this article. That is, the observability, favorability, or "operant/respondent" properties of the behavioral items could not explain the differences in predictability that were found between personality judgments from the self and knowledgeable others.

Further findings showed that the predictability of behavior was positively correlated with the operant versus respondent properties of the behavioral items. This finding holds for the predictability of behavior from self-judgments ($r = .55$, $p < .001$), and from acquaintances' judgments ($r = .54$, $p < .001$).

Finally, there was a tendency for more favorable items to manifest higher correlations between self-assessments and behavioral ratings from the videotapes ($r = .32$, $p < .05$). This finding (that self-assessments of favorable traits were more likely to have predictive validity than self-assessments of unfavorable traits) might suggest that people are more accurate at evaluating their own more favorable traits and is congruent with recent research by John and Robins (1994).

7. All analyses of correlations and differences between correlations reported in this subsection were performed on their z transforms.

DISCUSSION

Self and Others

Both personality and social psychology have traditionally assigned a prominent position to self-report as a source of data, and concomitantly often have emphasized the advantages the self enjoys over anyone else when evaluating one's own personality. The present results, however, suggest that the apparent informational advantages held by the self as a judge of immediate causes of behavior may be undermined by compensating disadvantages (see also Hofstee, 1994; Kenny, 1994). This conclusion is suggested by the finding that descriptions by a single close acquaintance had as great or better predictive validity than those provided by the self. In addition, the predictive validity of the aggregated judgments of just two acquaintances actually outperformed that of self-judgments in nearly every comparison in this study.

Kenny (1994) cautions that "in comparing the relative validity of self and peer, the researcher should be careful not to use peer reports that are averaged across multiple peers [because] . . . they would probably appear to be more valid than self-reports only because they are more reliable" (p. 193). Hofstee (1994), by contrast, views the fact that it is possible to gain judgments from multiple peers, but only one self, as the fundamental reason why it is "knowledgeable others" who are the best source for judgments of personality. He writes, "There is only one me, whereas there are many others who know enough about me to provide a more reliable average judgment. Thus, other things being equal, I am outnumbered and outperformed by the average other" (p. 153).[8]

Taking both points of view into account, it seems reasonable to conclude that a researcher should be careful not to interpret the superiority of aggregates of judges as applying to single judges. In the present data, the single judge was found to achieve only slightly if at all better pre-

8. It is important to note that Kenny and Hofstee might be interpreting the term "reliability" differently. Kenny could be referring to the fact that by using multiple peers, one would reduce the amount of error variance. This would make aggregated judgments appear more valid. However, this reduction of error variance may or may not be related to having more knowledge of the target. Hofstee, on the other hand, indicates that judgments are more reliable because multiple peers have more information to provide. We do not have adequate data to test these alternate possibilities. The necessary design would compare repeated self-assessments with judgments by multiple observers. However, we agree with Hofstee's view that because different judges have access to different information about the target, their combined judgments are more likely to be valid.

dictive validity than the self. But by the same token, the superiority of the aggregate of two judges in our data is surely more than a statistical artifact. It implies that the most reliable source of information about an individual's personality is to be found not in his or her self-judgments, but in the consensus of the judgments of the community of his or her peers. This difference in reliability would only increase, of course, as more informants (beyond the two in the present study) were recruited (Hofstee, 1994).

In any case, although fairly consistent, the differences in predictive validity between self and individual others found in this study are not by any standard large. Moreover, theoretical analysis as well as accumulated evidence would seem to indicate that actors might still be in a superior position for detecting the influence of immediate, situational factors on their behavior. Actors probably know better than anyone else the conscious intentions behind single, concrete actions, and in that way may be better at predicting what Hofstee (1994) calls "events."[9] But the present analyses and results may indicate that when dispositional regularities and thereby the prediction of more general aspects of behavior are the focus of interest, then observers enjoy their own advantages that more than compensate for those held by the self.

The Fish-and-Water Effect

One possible explanation for the present findings is that actors become so used to their own behavioral patterns that they become relatively unaware of them, whereas for observers, the differences between the behaviors of different actors in similar situations are highly salient. As McGill (1989) points out, "[I]n the actors' view, their traits hold constant while the situation and their behavior vary," making it difficult for actors to view their own dispositional consistencies (p. 191). "In the observers' view," McGill continues, "the situation holds constant while the behavior of different people in the situation varies" (see also Kahneman & Miller, 1986). It would then become difficult to detect one's own stable behavioral tendencies for roughly the same reason that fish are said to find it difficult to detect water. Observers, who can

9. However, Dianne Nilsen (personal communication, August 31, 1991) points out that actors' knowledge of their conscious intentions may not always constitute an advantage, because intentions do not always match subsequent behavior. As Nilsen puts it, "Perhaps the road to inaccuracy is paved with intentions."

more easily compare the person's unique attributes with those of others, would be better able to detect individual differences and dispositional regularities in behavior.

Perspective

The conclusion of the present research is also congruent with theoretical views that emphasize the ways in which one's personality is more visible from an external than from an internal perspective (e.g., Funder, 1991; Hofstee, 1994; Jones & Nisbett, 1971). Mead (1934) wrote, "The individual experiences himself as such not directly, but only indirectly . . . from the generalized standpoint of the social group as a whole to which he belongs" (p. 138). More recently, Bem (1972) has argued that "individuals come to 'know' their own attitudes, emotions, and other internal states partially by inferring them from observations of their own overt behavior. . . . [T]he individual is functionally in the same position as an outside observer" (p. 2). Even for the self, the outside observer's point of view is the only one from which personality is visible. This fact would seem to give the observer a natural advantage.

We strongly prefer this "perspective" explanation of the present findings to another possible explanation, which is that the excessive leniency of self-ratings, compared to acquaintances' ratings, leads them to be less accurate. As mentioned in the Results section, internal analysis of the present data found that the relative favorability of the behavioral items was uncorrelated with differences in predictive validity for self and acquaintance judgments. Moreover, in analyses presented elsewhere, we have found the enhancement tendency of the self over acquaintances to be smaller than the enhancement tendency that acquaintances have over strangers (Funder & Colvin, in press). On other occasions and in other data sets, we have found no self-enhancement tendency at all (Funder, 1980). In a particularly interesting study, Nilsen (1991) reported a decided advantage for the predictive validity of observers' over self-ratings, while at the same time finding no evidence for self-enhancement.

Direct versus Indirect Self-Report

The self-reports in the present study were direct (i.e., undisguised) self-descriptions. This fact raises the possibility that self-reports obtained through more complex questionnaire methods (e.g., those that include

correction scales), or perhaps tests with subtle items (e.g., the MMPI) that are not obviously related to the dispositions they are intended to measure, would yield superior behavioral predictions.

This issue will sound familiar to some readers. During the late 1960s and early 1970s, a large number of studies compared direct self-reports with a wide variety of what were called "indirect" methods, including questionnaires with disguised or subtle items, and projective techniques. The general conclusion of that research was that direct self-reports, of the sort used in the present research, performed as well as any of these alternatives (Hase & Goldberg, 1967; Mischel, 1972; Sherwood, 1966).

Interestingly, the usual validity criterion in that earlier research was observers' descriptions of personality. The basic finding was that direct self-reports predicted peers' judgments as well or better than did more indirect methods of assessment (McClelland, 1972). The present demonstration, by contrast, asked a new and very different question: What is the comparative validity of self- and peer judgments of personality, when the criterion is the prediction of psychologically meaningful behavior? The differences we found were nearly always in favor of the peers.

The I and the Me

William James (1892/1963) distinguished between the "I" of phenomenal experience and the "me" which is a person's more objective, public persona, and implied that the self might have privileged access to the I, but not to the me. James's distinction is relevant to the present research, which in effect addresses the me and not the I. The criterion for accuracy that we employed was the judgment's ability to predict overt behaviors of the sort our coders could extract from the observation of videotapes. As Kenny (1994) pointed out in his summary of other studies on this topic, "[T]he observers [who provide the criterion judgments] were, in a sense, peers" (p. 194). Judgments were used to predict meaningful behaviors as seen and evaluated (necessarily) from the outside; it perhaps should not be surprising (now that the results are in hand) that outside observers did as well or better than the self in predicting such outward appearances.

If the criteria in this research had been aspects of phenomenology or private emotional experience, the results might have been much different—the self might have enjoyed a profound advantage. On the other hand, if such a result were to be found, it would be open to a

parallel criticism. In such a case self-judgments would be superior to others' judgments in predicting self-judgments! A more fruitful resolution, therefore, might be to regard dispositional regularity in overt behavior as *one* fundamental manifestation of personality, and a criterion by which the advantage of others becomes more clearly apparent (Hofstee, 1994).

Alternative Criteria for Accuracy

Research on the accuracy of personality judgment needs to employ a diversity of criteria. In the present study, we evaluated the accuracy of personality judgments by their ability to predict independent, observational measures of the targets' overt behavior. We directly observed our targets in three experimental situations and coded their behavior at the level of analysis at which behaviors are intrinsically meaningful and externally observable, the common practice in modern behavioral assessment (Mischel, 1973). We believe our criterion is reasonable and defensible for the purposes of the present study. Moreover, it provides a good example of an attempt to use behavioral prediction as a criterion for accuracy. (For other, very recent examples, see Bank, Duncan, Patterson, & Reid, 1993; Borkenau & Liebler, 1993; Gangestad, Simpson, DiGeronimo & Biek, 1992; Jussim, 1993).

Still, we would want to be the first to acknowledge that our specific criterion demonstrates just one possible approach and that other research is needed. Particularly critical is the need for research that gathers behavioral assessments in other experimental as well as real-life contexts, and research that assesses behavior with other instruments besides the behavioral Q-sort used in the present study. When results from such diverse methods begin to converge—if they do—then we will be able to have much more confidence in our conclusions. As McCrae (1991) argues, "Much more can be done by comparing and contrasting different data sources" (p. 10). The present study represents one effort toward this end.

REFERENCES

Bakeman, R., & Gottman, J. M. (1986). *Observing interaction: An introduction to sequential analysis*. Cambridge: Cambridge University Press.

Bank, L., Duncan, T., Patterson, G. R., & Reid, J. (1993). Parent and teacher ratings in the assessment and prediction of antisocial and delinquent behaviors. *Journal of Personality*, **61**, 693–709.

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

Block, J. (1978). *The Q-sort method in personality assessment and psychiatric research.* Palo Alto, CA: Consulting Psychologists Press. (Original work published 1961)

Borkenau, P., & Liebler, A. (1993). 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.

Cairns, R. B., & Green, J. A. (1979). How to assess personality and social patterns: Observations or ratings? In R. B. Cairns (Ed.), *The analysis of social interactions: Methods, issues and illustrations* (pp. 209–226). Hillsdale, NJ: Erlbaum.

Coates, T. J., Killen, J. D., George, J., Marchini, J. E., Silverman, S., & Thorensen, C. (1982). Estimating sleep parameters: A multitrait-multimethod analysis. *Journal of Consulting and Clinical Psychology*, **50**, 345–352.

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

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

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.

Duval, S., & Wicklund, R. A. (1973). Effects of objective self-awareness on attribution of causality. *Journal of Experimental Social Psychology*, **9**, 17–31.

Fishbein, M., & Ajzen, I. (1974). Attitudes toward objects as predictors of single and multiple act criteria. *Psychological Review*, **81**, 59–74.

Fiske, D. W. (1979). Two worlds of psychological phenomena. *American Psychologist*, **34**, 733–739.

Frank, M. G., & Gilovich, T. (1989). Effect of memory perspective on retrospective causal attributions. *Journal of Personality and Social Psychology*, **57**, 399–403.

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

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

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

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, **102**, 652–670.

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 cross-situational 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 S. Briggs, W. Jones, & R. Hogan (Eds.), *Handbook of personality psychology*. San Diego: Academic Press.

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

Funder, D. C., & Harris, M. J. (1986). On the several facets of personality assessment: The case of social acuity. *Journal of Personality*, **54**, 528–550.

Funder, D. C., & Sneed, C. D. (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. (1993). Consensus, self-other agreement and accuracy in personality judgment: An introduction. *Journal of Personality*, **61**, 457–476.

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

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

Hase, H. D., & Goldberg, L. R. (1967). Comparative validity of different strategies of constructing personality inventory scales. *Psychological Bulletin*, **67**, 231–248.

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

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

Holzbach, R. L. (1978). Rater bias in performance ratings: Superior, self-, and peer ratings. *Journal of Applied Social Psychology*, **63**, 579–588.

James, W. (1963). *Psychology: A briefer course*. New York: Fawcett. (Original work published 1892)

John, O. P., & Robins, R. W. (1993). 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.

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

Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist*, **34**, 107–117.

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

Jussim, L. (1993). Accuracy in interpersonal expectations: A reflection-construction model of current and classic research. *Journal of Personality*, **61**, 637–668.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternative. *Psychological Review*, **93**, 136–153.

Kenny, D. A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review*, **98**, 155–163.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York: Guilford.

Kenrick, D. T., & Funder, D. C. (1988). Profiting from controversy: Lessons from the person-situation debate. *American Psychologist*, **43**, 23–34.

Langston, C. A., & Cantor, N. (1989). Social anxiety and social constraint: When making friends is hard. *Journal of Personality and Social Psychology*, **56**, 649–661.

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.

McClelland, D. C. (1972). Opinions predict opinions: So what else is new? *Journal of Consulting and Clinical Psychology*, **38**, 325–326.

McCrae, R. R. (1991, August). *The counterpoint of personality assessment: Self-reports and ratings*. Paper presented at the annual meeting of the American Psychological Association, San Francisco.

McCrae, R. R., Costa, P. T., Jr., & Busch, C. M. (1986). Evaluating comprehensiveness in personality systems: The California Q-set and the five-factor model. *Journal of Personality and Social Psychology*, **54**, 430–446.

McGill, A. L. (1989). Context effects in judgments of causation. *Journal of Personality and Social Psychology*, **57**, 189–200.

Mead, G. W. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.

Mischel, W. (1972). Direct versus indirect personality assessment: Evidence and implications. *Journal of Consulting and Clinical Psychology*, **38**, 319–324.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, **80**, 252–283.

Monson, T. C., & Snyder, M. (1977). Actors, observers and the attribution process: Toward a reconceptualization. *Journal of Experimental Social Psychology*, **13**, 89–111.

Moore, B. S., Sherrod, D. R., Liu, T. J., & Underwood, B. (1979). The dispositional shift in attribution over time. *Journal of Experimental Social Psychology*, **15**, 553–569.

Morrell, E. M., King, A. C., & Martin, J. E. (1986). The validity of cardiac inpatient self-report of smoking. *Behavioral Assessment*, **8**, 365–371.

Nilsen, D. (1991, August). *Understanding self vs. observer discrepancies in multi-rater assessment systems*. Paper presented at the annual meeting of the American Psychological Association, San Francisco.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego: Academic Press.

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

Peterson, C. (1980). Memory and the "dispositional shift." *Social Psychology Quarterly*, **43**, 372–380.

Roscoe, J. T. (1975). *Fundamental research statistics for the behavioral sciences* (2nd ed.). New York: Holt, Rinehart and Winston.

Scheier, M. F., Buss, A. H., & Buss, D. M. (1976). Self-consciousness, self-report of aggressiveness, and aggression. *Journal of Research in Personality*, **44**, 637–644.

Sherwood, J. J. (1966). Self-report and projective measures of achievement and affiliation. *Journal of Consulting Psychology*, **30**, 329–337.

Shrauger, J. S., & Osberg, T. M. (1981). The relative accuracy of self-predictions and judgments by others in psychological assessment. *Psychological Bulletin*, **90**, 322–351.

Sobell, L. C., & Sobell, M. B. (1975). Outpatient alcoholics give valid self-reports. *Journal of Nervous and Mental Disease*, **161**, 32–42.

Storms, M. D. (1973). Videotape and the attribution process: Reversing actors' and observers' points of view. *Journal of Personality and Social Psychology*, **27**, 154–164.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Century.

Watson, D., & Clark, L. A. (1991). Self- versus peer ratings of specific emotional traits: Evidence of convergent and discriminant validity. *Journal of Personality and Social Psychology*, **60**, 927–940.

Wiggins, J. (1973). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley.

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.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.