Perspectives on Personality:
The Relative Accuracy of Self Versus Others
for the Prediction of Emotion and Behavior

Jana S. Spain

High Point University

Leslie G. Eaton

David C. Funder

University of California, Riverside

ABSTRACT   This study compares the relative accuracy of targets' self-reports and other-reports of personality in predicting two criteria: (a) emotional experience in daily life and (b) behavior in the laboratory. Ratings of the targets' extraversion and neuroticism were obtained from two knowledgeable informants and the targets themselves. Target participants wore an electronic signaling device ("beeper") for eight days and rated positive and negative emotions at four randomly selected times each day. The participants also interacted with an opposite-sex stranger in a laboratory context and their behavior was coded from videotapes. Targets' self-reports of personality were consistently more

accurate than other-reports in predicting daily emotional experience. Self-reports also outperformed other-reports in predicting extraversion-related laboratory behaviors, but not neuroticism-related behaviors. The relative accuracy of self- and other-reports of personality would seem to depend on the criterion employed; self-reports are clearly better for the prediction of emotional experience, while for behavior the picture is mixed.

Over the past decade, a large body of empirical research has clearly demonstrated that laypersons' judgments of the personality traits of both friends and strangers are generally accurate (Funder & West, 1993). Consequently, researchers have now turned their attention from demonstrating whether personality judgments are accurate to identifying the conditions under which judgments are more or less likely to be accurate. One factor that could affect the accuracy of a personality judgment concerns the identity of the judge: the target person or others.[1] Although both self- and others' judgments have been used to assess personality, within personality and social psychology targets' self-judgments have long been presumed to be more accurate (Wiggins, 1973). In fact, self-judgments are often the standard by which the validity of other measures is gauged (e.g., Paunonen, 1989). After all, the logic goes, who should be in the best position to judge personality: the target individual or someone else? The field's overwhelming reliance on data from self-report questionnaires suggests that most researchers assume that the answer to this question is "The target, of course" (Craik, 1986; Moskowitz, 1986).

That assumption was called into question, however, by a recent study by Kolar, Funder, and Colvin (1996), which compared the predictive validity of self and others' reports. Kolar et al. demonstrated that, when the criterion for accuracy was behavioral prediction, judgments by close acquaintances often predicted behavior as well as did self-judgments. In fact, the composite (averaged) judgment of just two other acquaintances generally outperformed self-judgments of personality. Kolar et al. argued that the differences in perspective between others and the target individual might have contributed to this outcome. They maintained that because others' attention is focused on the target person, while the target is focused on his or her immediate environment, others might be in a better

---

1. The terms "target(s)" and "target individual" are used throughout this article to refer to the participants of the study or the individual whose personality is being judged.

position to observe individual differences or patterns in behavior. As a result, behavioral manifestations of personality might be more easily judged from the outside (Bem, 1972; Funder, 1991; Hofstee, 1994; Jones & Nisbett, 1971).

Although others' personality judgments might be more accurate for predicting overt behaviors (e.g., is talkative, initiates humor, acts condescending), it seems possible that any such advantage might not be evident if alternative criteria for accuracy were employed. For example, if the criterion for accuracy were to be the prediction of aspects of personality that are more internal in nature, self-judgments might have the advantage. In particular, emotional experience is one manifestation of personality that might be more easily judged from an internal rather than an external perspective. Others might occasionally have a better view for overt behavior, but for everyday emotional experience, the target individual would seem to have the best seat in the house.

In order to test this possibility and extend our knowledge of the conditions under which personality judgments from different sources are more likely to be accurate, in this study we compared the relative accuracy of self-report and other-report by examining their utility in predicting two criteria for accuracy: (a) the prediction of emotional experience in daily life and (b) the prediction of overt behavior in laboratory situations (in a quasi-replication of Kolar, Funder, & Colvin, 1996).

Emotion and Behavior as Criteria for Accuracy

Determining the accuracy of a personality judgment by testing its ability to predict theoretically relevant criteria is not a new approach (Funder & Colvin, 1991; John & Robins, 1994). Using this approach to determine the relative accuracy of self and others' judgments, however, is somewhat new and can present challenges for researchers (Kenny, 1994; Kolar et al., 1996). These challenges arise from the necessity to find criteria that are independent of self- and other-judgments of personality, the validity of which are being compared. For the present study, we sought independent assessments of both emotional experience and behavior and focused our attention on two personality characteristics that are related to both emotion and behavior. For our personality factors, we chose to focus on Extraversion and Neuroticism, two major factors of the Five-Factor Model (McCrae & Costa, 1985; Norman, 1963). Dimensions with

fairly similar content to these factors are found in other theoretical conceptualizations and it is reasonable to expect both of these dimensions of personality to be at least potentially relevant to a wide range of social behaviors as well as to emotion.

The link between personality and emotion. If the prediction of emotion is to serve as an appropriate criterion for the accuracy of personality judgments, it is necessary to consider carefully how emotion might be related to personality. In recent years, two largely independent dimensions of emotion, negative and positive affect, have consistently emerged from analyses of mood adjectives as well as self-rated mood (see Watson & Tellegen, 1985, for a summary). Negative affect is characterized by feelings of distress including anxiety, anger, fear, disgust, and sadness, whereas positive affect is characterized by pleasant feelings such as excitement, happiness, joy, and interest. These emotions are evident in facial and vocal emotional expressions and numerous studies suggest that they can be detected and correctly recognized by human judges (Ekman, 1973; Ekman, Friesen, & Ancoli, 1980; Ekman, Friesen, O'Sullivan, & Scherer, 1980). And, regularities in the expression of personality characteristics through these emotions are evident by the time an individual reaches adulthood (Magai & Hunziker, 1993).

Although other personality traits are relevant to these emotion dimensions (McCrae & Costa, 1991; Watson & Clark, 1992), the broad personality factors of Neuroticism and Extraversion have been linked to both the experience and the expression of these emotions in adulthood most consistently. Consequently, we chose to focus our attention on these two personality factors. Theoretically, Costa and McCrae (1980, 1984) have maintained that the specific facets of Neuroticism probably share a common origin in negative affect, while aspects of Extraversion have a common origin in positive emotionality. Empirically, the results of several studies support the idea that the broad factors of Neuroticism and Extraversion form the basic personality structure underlying emotion (Emmons & Diener, 1985; Meyer & Shack, 1989; Watson & Clark, 1992).

The link between neuroticism and negative emotions is well documented. Numerous studies have demonstrated relationships between neuroticism and various measures of negative emotionality (McCrae & Costa, 1991; Watson & Clark, 1984). Negative emotion has been proposed to be enhanced and maintained through the relationship between

neuroticism and the increased incidence of retrieval of negative personal memories, negative cognitions and attributions, and heightened stress reactivity (Ormel & Wohlfarth, 1991; Ruiz-Caballero & Bermudez, 1995). The relationship between neuroticism and negative emotion is also evident in nonclinical samples and appears to be stable over time and across situations (Gross & John, 1995; Izard, Libero, Putnam, & Haynes, 1993; McLennan, Buchanan, & Bates, 1994). For instance, in laboratory experiments neurotics have been shown to be more reactive to negative mood induction procedures (Larsen & Ketelaar, 1989), while in experience sampling studies neurotics tend to report experiencing more frequent and more extreme negative emotions in daily life (Bolger & Schilling, 1991).

Positive emotion is central to theories of extraversion and the relationship between extraversion and positive emotions has also been documented, although not as extensively (Costa & McCrae, 1980; Watson & Clark, 1992; Watson, Clark, McIntyre, & Hamaker, 1992). Links similar to those identified between neuroticism and negative moods have been identified between extraversion and the experience of positive moods. For example, extraverts have been shown to possess a heightened reactivity to positive mood induction and tend to report experiencing positive affect more frequently in daily life (Diener & Larsen, 1993; Larsen & Ketelaar, 1989, 1991; Watson et al., 1992).

Measuring emotional experience.   Despite the highly developed technology for measurement of other biologically related phenomena, a methodology for the objective measurement of emotional experience is, to our knowledge, nonexistent. Using various physiological measures (e.g., galvanic skin response; facial electromyographic and electrocortical activity), researchers currently can measure whether or not a subject is experiencing an emotion, and make rough inferences regarding the intensity of that emotion. No technology, however, can directly determine *which* emotion a subject is experiencing. Although some major steps toward mapping specific biological cues to discrete emotions have been made, the maps researchers have developed to date rely upon linking regularities in the self-report of emotional experience to measured biological states. The identified links may or may not indicate a one-to-one correspondence between the biological state and the reported emotional experience (see Cacioppo, Klein, Berntson, & Hatfield, 1993, for a review).

It is probably impossible to obtain a measure of emotion that is completely independent of self-reports. A reasonable alternative for the purposes of this study, however, is to use a measure of emotional experience that is independent of personality judgments. Experience sampling "beeper" methods, which involve signal contingent reporting of behavior or emotions, offer one such option (Spain, 1994; Wheeler & Reis, 1991). Participants report their emotional state when signaled by the researcher. Signaling can occur at randomly selected times and reporting can be done immediately following the emotion or behavior of interest, thereby reducing many of the difficulties associated with other types of self-reports (Csikzentmihalyi & Larson, 1987; Hormuth, 1986). If these methods are used, the measurement of emotional experience is separated from the self-reported personality judgment because the measurement of emotion takes place across multiple times in everyday settings and occurs at times determined by the researcher rather than the participant. Experience sampling measures of emotions approximate "online" measures of emotional experience in daily life, providing a viable alternative to traditional, retrospective summary measures of emotion.

Measuring behavior. The measurement of behavior can also raise some complex issues. The direct measurement of personality-relevant behavior is an expensive and time-consuming enterprise, which may account for its relative rarity in personality research (Funder, Furr, & Colvin, 1999). The researcher must either find or set up contexts in which behavior can be directly observed, choose an appropriate language and technique for describing and coding this behavior, and then actually perform the coding in a reliable and valid manner. In the present research, as in other studies, each of these steps required a major effort.

First, as will be described in more detail in the Method section, we put approximately 180 undergraduates through an experimental procedure in which they engaged in an unstructured interaction with a peer of the opposite sex. The setting was designed to allow a free range for various social behaviors, while also being interesting and personally involving for the participants. Second, we developed the Riverside Behavioral Q-sort (RBQ), a technique for conceptualizing and recording a wide variety of behaviors described at a level of analysis appropriate for comparison with personality ratings (Funder, Furr, & Colvin, 1999). Third, a large number of trained coders used the RBQ to describe the behavior of the participants in the experimental context, a process that

required years to complete (longer than was required to run the experi-mental contexts themselves). Then, specific behaviors relevant to the personality factors being utilized in this study were identified.

This brief summary can only hint at the many specific choices that had to be made and the procedures that had to be developed to allow the construction of a set of behavioral criteria for the research reported in this study. More detail can be found both in the Method section and also in articles describing the development of the RBQ (Funder, Furr, & Colvin, 1999) and the development of experimental contexts for the observation of personality-relevant behavior (Funder & Colvin, 1991; Kolar, Funder, & Colvin, 1996).

Perspectives and Accuracy

Who has the best perspective from which to judge personality? The answer to this question may not be as simple or as clear as the field seems to have presumed. As suggested above, it is possible that the answer to this question may depend upon which criterion is used for accuracy and which traits are examined. According to the Realistic Accuracy Model (RAM; Funder, 1995), accurate personality judgment is dependent upon attributes of the target of judgment, the trait being judged, the behavioral cues relevant to the trait, and the nature and ability of the judge. Beginning with the "neo-Allportian" assumption that personality traits are real attributes of individuals (Funder, 1991), the model asserts that judg-mental accuracy derives from the relevance, availability, detection, and utilization of behavioral cues to personality traits.

The target and others may use some of the same cues for their judgments of personality. They may have similar understandings of the relevance of the cues for a particular characteristic. Nonetheless, the target individual and others sit in different places, see different things, and have access to varying amounts and different kinds of information. Because these judges of personality have different perspectives, one should expect their judgments to differ in accuracy to the extent they have access to relevant behavioral cues to the characteristic being judged.

Prediction of emotional experience: The view from the inside.   It would seem that patterns of emotional experience should be more easily observed from a closer perspective, leading targets' self-judgments to be more valid. In daily life, the experience of emotion is by and large a private affair.

Individuals can easily be experiencing positive and negative emotions of varying intensities and keep those experiences private (Buck, 1984; Izard, 1991). In fact, it is often socially unacceptable to display emotional experience in public, particularly when the emotion is negative (Saarni, 1993). By adulthood, individuals have become quite good at hiding their emotional experiences by controlling facial and other upper body expressions (Ekman, 1992; Ekman & Friesen, 1975, 1980). The private nature of emotional experience could be expected to limit the access to relevant cues for others.

Still, whether the target individual will truly have an advantage in the prediction of emotional experience is open to question. For example, the person's focus on the external environment could impair the correct detection and labeling of emotion. When individuals are in doubt about the source of the physiological arousal that accompanies their emotions, they can look to the external world to help them identify and label emotions (Bem, 1972; Schacter & Singer, 1962). At times, individuals may become so focused on the external environment that they misattribute their arousal and incorrectly judge their own emotions (Dutton & Aron, 1974; Valins, 1966). Because others also have access to this external environment, their judgments about emotional experiences could potentially have comparable validity.

Moreover, despite their ability to control some emotional expressions, individuals may not always monitor or successfully manage to contain expressions, particularly with others who know them well. Although it is likely that individuals frequently attempt to control outward expressions of negative emotionality, they may not always be successful (Buck, 1984). For example, when persons are experiencing high-arousal negative emotions such as fear, anger, or anxiety, cues may leak out through nonverbal channels that are unattended and uncontrolled.

It is likely that the accuracy of others' judgments might be related to the type of emotion that is used as the criterion. Positive and negative emotions differ in social desirability, degree of spontaneous expressivity, number of facial and bodily cues, and frequency of occurrence in daily life (Buck, 1984; Lewis & Haviland, 1993). Consequently, the amount of information available that is relevant to the judgment of traits related to these emotions also may vary considerably. Because positive emotions are generally seen as socially desirable and are frequently expressed spontaneously, individuals may not even feel a need to monitor or control their outward expression, and so others may be given relatively free

access to positive emotion cues (Buck, 1984; Ekman & Friesen, 1975). Negative emotions are viewed as less socially desirable, they may be less likely to be expressed, and they tend occur less frequently than positive emotions. Consequently, the judgment of negative emotions may tend to be quite difficult, particularly for judges who are relying on overt manifestations of the emotion for their judgment.

The prediction of overt behavior: The view from the outside. It was perhaps surprising to some that Kolar et al. (1996) found others' personality judgments to be equally as accurate as and sometimes more accurate than targets' self-judgments for predicting behavior. There are a number of reasons to expect that, even for behavioral criteria, the target individual should have the clearest view when it comes to judging one's own characteristics. For one, self-judgments are based on the largest quantity of information. Accurate judgments of personality should require information about cross-situational consistency and temporal stability of both public and private behaviors. Therefore, the judge with the greatest access to this kind of information, the target, should be in a relatively privileged position.

Kolar et al. (1996), however, argued that there is only one target person, while there is an infinite number of other judges who can provide judgments (cf. Hofstee, 1994). One purported advantage of the use of others' judgments is that researchers can combine the judgments of multiple others, leading to a personality measure with the opportunity to be more reliable than a single self-judgment could be. If these others also observe the target individual's behavior in a variety of situations, the potential increase in the amount of information alone could result in their aggregated judgments being at least as accurate as, if not more accurate than, self-judgments.

The target individual could be at a disadvantage in other ways. Self-protective biases might prevent one from correctly assessing his or her own characteristics (Kenny, 1994). Individuals may be less inclined to admit that they possess undesirable traits, describing themselves using only those traits considered socially desirable in our culture. This tendency may be the result of impression management or result from ego defensiveness. And, although this tendency may serve an important protective function for one's ego and self-esteem, it may contribute to inaccuracy. An individual's moods can impact the reliability and accuracy of self-judgments and even one's own personality traits might actually

prevent one from providing self-descriptions that are accurate (Bradley, Mogg, Perrett, & Galbraith, 1993; Hjelle & Bernard, 1994; John & Robins, 1994). For example, John and Robins (1994) found that the tendency for narcissists to self-enhance resulted in their self-descriptions having less predictive validity than the judgments of other individuals.

And, finally, it is possible that an individual might be less able than others to provide an accurate judgment of his or her traits, not because of a motivation to see oneself as better (or worse) than one really is, but because he or she may be in relatively poorer position to view his or her own behavioral consistencies. An individual is constantly responding to situational changes, deciding what to do and say next, and altering behavior to fit new circumstances (Jones & Nisbett, 1971). What is salient and visible to the person is the situation, whereas what is salient and visible to others is the person's behavior (Frank & Gilovich, 1989; Storms, 1973). As a result, it is possible that the person may be very aware of the variability in his or her behavior but may be less aware of the consistency evident in his or her behavioral patterns than are others. These consistent and individually distinctive patterns of behavior might be more easily observed from the more distant perspective of an observer.

The Relative Accuracy of Self Versus Others' Judgments

In this study, we compared the relative accuracy of self and others' personality judgments using the criteria of emotional and behavioral prediction. We compared the ability of targets' self-judgments of neuroticism and extraversion to predict their aggregated daily positive and negative emotional experiences (measured through experience sampling methods) and relevant behaviors (coded from a videotaped interaction) to the judgments of others who knew the targets well. When the criterion was the prediction of everyday emotional experience, we predicted that the target would have a clear advantage. Because of the unique access of the target person to information about emotion, we expected this advantage to hold even when a single self-judgment was compared to the composite judgments of two others. We expected, however, to replicate the general pattern of findings of Kolar et al. (1996) for the judgment of overt behaviors. Specifically, when the criterion for accuracy was

behavioral prediction, we expected others to do equally well and some-
times better than the target person.

METHOD

The data for this study were collected as part of the Riverside Accuracy Project,
a large research project on the accuracy of personality judgment. During a
3-to-4-month period, 182 (91 male and 91 female) targets completed numerous
self-report measures, were videotaped interacting with acquaintances and
strangers on several occasions, reported on their daily experiences, and recruited
several knowledgeable informants who described their personalities. Although
various aspects of these data have been reported elsewhere (e.g., Creed &
Funder, 1998a; Creed & Funder, 1998b; Funder, Kolar, & Blackman, 1995; Furr
& Funder, 1998; Sneed, McCrae, & Funder, 1998) the present analyses are new
and the data pertaining to emotional experience have not been published
previously.[2] It also can be noted that the earlier, relevant study by Kolar et al.
(1996) was based on an independent sample of participants at a different
university.

Participants

The Riverside Accuracy Project target participants were undergraduate students
at the University of California, Riverside, recruited through flyers posted
campuswide and announcements made in various classes. Each target sub-
sequently recruited two college acquaintances to serve as informants to describe

2. Funder, Kolar, and Blackman (1995) examined acquaintanceship, similarity, commu-
nication, overlap, and mutual accuracy as the bases for interjudge agreement. Furr and
Funder (1998) examined correlates of personal negativity. Creed and Funder (1998a)
examined personality and behavioral correlates of private self-consciousness. Creed and
Funder (1998b) focused on the construct of social anxiety, identifying both the person-
ality characteristics and interaction styles of socially anxious individuals. Sneed, McCrae,
and Funder (1998) examined laypersons' understanding of the Five-Factor Model and
the behavioral manifestations of each factor. Study 2 of Sneed et al. included analyses
similar to those conducted here to compare laypersons' diagnosticity ratings of behaviors
for each factor and the observed correlations between those behaviors and the factors.
The behavior ratings used in the Sneed et al. paper, however, were composite ratings
aggregated across an unstructured, a cooperative, and a competitive interaction. Our
analyses include only the RBQ ratings from the unstructured interaction, two of the five
factors, a subset of factor-relevant behavioral items selected using an independent
method, and those participants with complete data for all of the variables examined here.
Consequently, the correlation coefficients reported in Sneed et al. are different from those
reported here. The beeper emotion variables have not been used in any other study.

the target's personality.[3] All participants received compensation for their participation. From this initial pool, we selected 82 targets with complete data on all measures for the present study.

Personality Judgments

Judgments of the targets' neuroticism and extraversion were obtained from the target and his or her two college acquaintances using the NEO Personality Inventory (NEO-PI; Costa & McCrae, 1985).[4] The NEO-PI is a 181-item questionnaire developed to measure the personality dimensions in the Five-Factor Model of personality. The NEO-PI was selected because it has demonstrated impressive reliability and validity for both self and others' ratings and provides assessment of broad personality dimensions. Targets and their college acquaintance informants were administered the NEO-PI in the laboratory.

Emotion Measures

Target participants wore an electronic signaling device ("beeper") for eight days. Each day, between the hours of 9:00 A.M. and 8:00 P.M., participants were signaled or "beeped" at four different times. To determine signal times, the time period was divided into four intervals and one time was randomly selected from each interval. Upon being signaled, each target completed a brief questionnaire describing his or her current emotions and activities. This questionnaire was adapted from other experience sampling studies (e.g., Csikzentmihalyi & Larson, 1984).

For the emotion ratings, targets were asked to rate how strongly they were feeling each of 12 emotions just before their beeper sounded. Response options ranged from 0 (*not at all*) to 4 (*extremely*). Each emotion score was the average rating of the emotion averaged across all of the target's beeper responses. Negative emotions included "nervous," "scared," "sad," "hostile," "distressed," and "upset," and positive emotions included "happy," "joyful," "pleased," "energetic," "interested," and "calm."

---

3. Participants were asked to recruit two college friends who knew them well. When these informants began their participation in the study, they were arbitrarily designated as either Informant No. 1 or Informant No. 2 by our research assistants.

4. All analyses focus exclusively on the factors of Extraversion and Neuroticism, because only these two factors have been theoretically and empirically linked to both emotion and behavior and, consequently, our hypotheses involve only these factors. Results for the entire Five-Factor Model are available upon request from the first author or from the Riverside Accuracy Project home page at http://www.psych.ucr.edu/faculty/funder/RAP/perspweb.htm.

Participants were requested to fill out the questionnaire as soon as possible after being signaled. In order to encourage compliance and reduce delayed reporting, participants were asked to turn in each completed questionnaire the following morning. Compliance with the experience sampling procedure was quite good. The average subject completed 86% of the 32 beeper questionnaires within 30 minutes of being signaled. This response rate is similar to that obtained by other researchers using experience sampling methods including Csikzentmihalyi and Larson (1984) and Cantor et al. (1991). Information provided on incomplete beeper reports as well as informal interviews with participants following the beeper portion of the study indicated that missing reports were caused by a number of factors. One source of missing data was the failure of beepers to signal participants at the scheduled times. Although some of these failures were due to mechanical problems (e.g., low batteries, computer or relay station malfunctions) most were attributable to operator difficulties. For example, participants were instructed to clear their beeper immediately after each signal, thereby preparing it to receive the next signal. At times, subjects did not properly clear their beepers and, as a result, a number of their reports went missing. Participants had been instructed to inform the researchers if their beeper did not function properly and so many of these difficulties were quickly identified and remedied.

Although participants were encouraged to wear their beepers at all times, they were allowed turn their beeper off when they did not want to be interrupted. A few participants did this, usually during exams or when they were sleeping. It does not appear, however, that our participants chose to selectively respond to signals. In fact, most of the times that beepers were not worn continuously were the result of accidentally leaving the beeper at home rather than any conscious decision to not carry it.

A final source of missing data was due to delays in the response to signals. One of the distinctive strengths of beeper methods is that they result in behavioral reports that are close to "online" measure of behavior, thereby reducing bias that may result from retrospective reporting. Therefore, we decided that late reports would not be included in the present analyses. Participants were asked to record both the time the beeper sounded and the time they completed the report. These reported times were then matched against the researchers' records of actual signal times. Nonmatching times and reports that were completed more than 15 minutes after the person was signaled were counted as missing reports and those emotion ratings were not included in the analyses reported here. Using this cutoff, we retained 71% of the original data and participants completed an average of 20 beeper reports.

Behavioral Measures

Previously unacquainted pairs of target participants, one male and one female, were scheduled by telephone to arrive for their initial experimental session. Upon arrival in the laboratory, it was ensured that the two participants had never met. They were then led into the laboratory and were seated on a couch in front of a video camera that was not concealed. The participants were encouraged to "talk about whatever you'd like" by the researcher. The researcher activated the camera and then departed to return 5 minutes later. This unstructured interaction was designed to create a type of commonly experienced interpersonal situation, allowing aspects of the participants' personality to be expressed in their behavior.

Coding of behavioral interactions. Targets' behavior in this session was independently coded by four trained research assistants using the 64-item RBQ, a revision of the 62-item Behavioral Q-sort used by Kolar et al. (1996) and other research related to the Accuracy Project (e.g., Funder & Colvin, 1991). The RBQ was designed to provide information about behavior at a psychologically meaningful level of analysis (Funder, Furr, & Colvin, 1999). Therefore, each of the items in the RBQ describes personality-relevant social behavior at a midrange of generality between microlevel behaviors (e.g., frequency of hand gestures) and more macrolevel impressionistic behaviors (e.g., is successful). (For a discussion of levels of analysis in behavioral coding, see Funder & Colvin, 1991; Funder et al., 1999.) These behaviors, such as "acts irritated" or "expresses warmth," were rated by the coders, who sorted the 64 RBQ items into a nine-category, forced choice, quasi-normal distribution ranging from *not at all descriptive* (1) to *highly descriptive* (9) of the participant's behavior. In this way, each behavioral item received a rating ranging between 1 and 9 from each of the coders and the ratings for each participant were averaged across coders to obtain a score for each RBQ item.

To ensure reliability for the behavioral coding, each coder's ratings were compared with every other set of ratings for the same participant. Coders' ratings were used in the overall composite score only if they correlated at least .30 with those of two other coders and at least .25 with the third coder. A coding that failed to achieve this threshold was repeated or replaced. This procedure ensured a minimum alpha reliability of .60 for each composite behavioral coding; the actual alpha reliability of the composite profile (across all 64 items) for this session was .81. For the present analyses, more germane than the profile reliability is the reliability of the coding of each item. The reliability estimates of the ratings for each of the 64 items, calculated across the four coders, ranged from .80 to .09[5] with a median of .55.

---

5. Only four items had reliabilities below .30. The content of these items usually referred to behavior that was seldom expressed in the situation studied here (e.g. "Expresses

Selection of personality-relevant behaviors. In order to select specific behaviors that would be theoretically relevant to our personality factors, we relied on the results of a factor analysis reported in McCrae, Costa, and Busch (1986). To test the comprehensiveness of the Five-Factor Model, McCrae et al. factor-analyzed ratings from the California Q-set (CQ; Block, 1978), a frequently used personality assessment device, and identified CQ items that defined each of the five factors. Twenty-two CQ items defined the Extraversion factor and 31 items defined the Neuroticism factor in that study.

Thirty-seven of our 64 RBQ items were originally written to directly correspond to descriptive personality statements included in the CQ. For example, an item in the CQ reads "Has social poise and presence; appears socially at ease" and the corresponding RBQ item is "Exhibits social skills." Another CQ item is "Is basically anxious" and the RBQ item is "Shows physical signs of tension and anxiety." Thirteen of these RBQ items correspond to CQ items that defined the Extraversion factor and 17 items correspond to CQ items that defined Neuroticism in the McCrae et al. (1986) study. We selected these cognates as our specific extraversion-relevant and neuroticism-relevant behaviors.

RESULTS

We carried out a series of analyses to test each of our hypotheses, first comparing targets' self-judgments to each individual informant's judgment and then comparing the self-judgment to the composite judgment of both informants. We decided that our primary analyses would examine each of the emotions and behaviors individually. Although single items are generally less reliable than an aggregated score, which tends to attenuate the resulting correlation coefficients, the analysis of individual items is more consistent with the procedures used in Kolar et al. (1996). Following the analysis of individual emotions and behaviors, we created composite emotion and behavior variables, as described below, and then conducted similar analyses with the composite variables.

For emotions, we expected to find that the targets' self-judgments would be more accurate than those of others. To test this hypothesis, we first correlated the targets' average rating for each of the six positive emotions with the self and each individual informant's NEO-PI Extraversion scores (see Table 1). As can be seen in the table, self-reports of personality

interest in fantasy or daydreams"). Several other items received the same ratings across all targets, thereby restricting their range. See Funder, Furr, and Colvin (1999) for a more detailed discussion of issues related to item reliability for the RBQ.

consistently yielded larger correlations than other-reports. Of greater interest, however, is the average size of the correlations across all the positive emotions. Following a procedure similar to that used by Kolar et al. (1996), the specific positive emotion correlations were averaged[6] to obtain a mean correlation size for each type of judge. Correlation coefficients in the opposite direction from that expected theoretically were treated as negative numbers for the averaging of the correlations. The resulting averages are shown at the bottom of the table. Paired comparison *t*-tests were conducted between the mean self-values and the mean values for the informants. Means shown in boldface are significantly different from the self-mean at $p < .05$, one-tailed. When the criterion was the prediction of everyday positive emotions, the average for the self (mean = .34) was significantly higher than the average for each individual informant (mean = .19 and .15 for Informant No. 1 and Informant No. 2, respectively). The same series of analyses was performed using self- and other-reports of Neuroticism and the target's negative emotion ratings. The same pattern was evident in the relationships between negative emotions and Neuroticism judgments (see Table 2). The average correlation size was .31 for the self-judgment but only .07 and .06 for the informants' judgments.

One purported advantage that others' judgments have over self-judgments, however, is that although there is only one person who can provide a self-judgment, you can obtain ratings from multiple others to form an aggregated composite judgment that is likely more reliable than any single score provided by the self (Kenny, 1994). Because Kolar et al. (1996) had found that the composite of just two peers was frequently more accurate than the single self-judgment, we wanted to determine if any advantage for others would be evident when the target was compared to multiple others. First, we formed a composite Extraversion judgment for the two informants averaging their NEO Extraversion ratings and a composite Neuroticism judgment averaging their NEO Neuroticism ratings.[7] These composite informant personality judgments were then correlated with the targets' positive and negative emotion scores, an average correlation value was calculated, and the averages for self and the informant composite were compared using the same procedures

---

6. All averaging of correlations employed Fisher's *r* to *z* transformation method.

7. The interjudge correlations for Informant No. 1 and Informant No. 2 were .56 for Extraversion and .39 for Neuroticism.

Table 1
NEO Extraversion Correlates of Positive Emotion Ratings for Self,
Individual Informants, and Informant Composite (*N* = 82)

		Source of Personality Description		
Emotion	Self	Informant #1	Informant #2	Informant Composite
Calm	.14	−.02	.01	−.00
Energetic	.38*	.13	.18	.18
Happy	.43*	.33*	.32*	.36*
Interested	.28*	.16	−.03	.07
Joyful	.33*	.21	.28*	.28*
Pleased	.38*	.28*	.15	.24
Average	.34	**.19**	**.15**	**.19**
Positive Emotion Composite	.40*	**.23***	**.18**	**.23***

Note. Average and composite informant values shown in boldface are significantly lower than the average self value at *p* < .05, one-tailed. All averaging of correlation values employed Fisher's *r* to *z* conversion.
* *p* < .05.

Table 2
NEO Neuroticism Correlates of Negative Emotion Ratings for Self,
Individual Informants, and Informant Composite (*N* = 82)

		Source of Personality Description		
Emotion	Self	Informant #1	Informant #2	Informant Composite
Distressed	.29*	.04	.02	.04
Hostile	.27*	−.02	−.00	−.01
Nervous	.32*	.07	.03	.06
Sad	.34*	.19	.19	.23*
Scared	.24*	.15	.12	.16
Upset	.29*	.02	−.02	.00
Average	.31	**.07**	**.06**	**.08**
Negative Emotion Composite	.39*	**.09**	**.07**	**.07**

Note. Average and composite informant values shown in boldface are significantly lower than the average self value at *p* < .05, one-tailed. All averaging of correlation values employed Fisher's *r* to *z* conversion.
* *p* < .05.

described above. The results of these analyses are listed in the fourth columns of Tables 1 and 2. The mean values for the targets' self-judgments were still significantly higher than the mean values for the informant composite judgments. Even the combined judgment of two knowledgeable informants did not perform as well as a single self-judgment.

Next, we created composite positive emotion and negative emotion variables. To create these composites, targets' emotion ratings were first transformed to z-scores and then summed. The alpha reliability for these positive emotion and negative emotion composites was .89 and .84, respectively. These scores were then correlated with self and others' personality judgments. The resulting correlation coefficients are listed in the bottom row of Tables 1 and 2. The results of t-tests for dependent correlations (Cohen & Cohen, 1983) indicated that the correlations for the self-judgments were uniformly larger than the correlations for others' ($p < .05$, one-tailed). The correlation between the self-judgment of Extraversion and the positive emotion composite was .40 and the correlations for others ranged between .18 and .23. The corresponding coefficients for Neuroticism judgments and the negative emotion composites were .39 for targets' self-judgments but ranged between .07 and .09 for others' judgments. Thus, when the criterion for accuracy was emotion in daily life, targets were consistently more accurate than were others.

But what happens when the criterion for accuracy is the prediction of behavior? We expected that when the criterion for accuracy was behavioral prediction, others' judgments would do equally as well as and sometimes better than the targets' self-judgments. In order to test this hypothesis, we correlated each of the targets' 13 extraversion-relevant RBQ behaviors with both self and others' NEO Extraversion ratings. The resulting coefficients are, of course, relatively small in magnitude, as one would expect to find with single item behavioral ratings of one 5-minute interaction. Table 3 lists the individual Pearson product–moment item correlations for self and each individual informant. As can be seen in the table, the targets' self-judgment usually resulted in higher item correlation coefficients than others' but occasionally others' judgments resulted in higher item correlations than the self-judgment. More important, however, is the average size of the correlations across the 13 extraversion-relevant behaviors. After reversing items where theoretically appropriate, the individual correlations were averaged. Once again, correlation coefficients in the opposite direction from that expected theoretically were treated as negative numbers for the averaging. For extraversion-relevant

Table 3
NEO Extraversion Correlates of Riverside Behavioral Q-Sort (RBQ)
Items for Self, Individual Informants, and Informant Composite
(N = 82)

RBQ Item	Self	Informant #1	Informant #2	Informant Composite
		Source of Personality Description		
Exhibits social skills	.37*	.07	−.01	.03
Is reserved and unexpressive (R)	−.26*	−.14	−.01	−.09
Is talkative	.18	.08	.01	.05
Initiates humor	.15	.10	−.05	.02
Seems likable	.13	.06	.17	.13
Expresses warmth	.06	.05	.11	.09
Is expressive in face, voice or gestures	.19	.15	.02	.09
Keeps partner at a distance (R)	−.21	−.10	−.18	−.16
Says or does interesting things	.15	.07	.13	.12
Gives up when faced with obstacles (R)	.01	.14	−.09	.02
Offers advice	−.09	−.07	−.02	−.05
Speaks fluently and expresses ideas well	.21	.05	.01	.03
Partner seeks advice from subject	.07	.02	.03	.02
Average	.15	**.05**	**.05**	**.06**
Extraversion Behavior Composite	.36*	**.13**	**.13**	**.15**

Note. Average and composite informant values shown in boldface are significantly different from the average self value at $p < .05$, two-tailed. All averaging of correlation values employed Fisher's r to z conversion.
(R): Sign of correlation reversed for averaging.
* $p < .05$.

behaviors, the average self value (mean = .15) was significantly higher ($p < .05$, two-tailed) than the value for either of the individual informants (mean for each informant = .05).

Does this advantage for the target individual disappear when his or her judgment is pitted against the composite of two or more judges? In order to answer this question, the self-judgments were again compared to the composite informant judgments. Pearson product–moment correlation coefficients were calculated between each RBQ item rating and the composite informants' Extraversion judgments and mean values were calculated. The results are shown in the last column of Table 3. Contrary to the findings of Kolar et al. (1996), we did not find that the composite of the two informants was more accurate than the targets' self-judgment. In fact, the mean value for the self-judgment was still significantly higher than the mean value for the informant composite judgment (mean = .06).

The same set of analyses was performed using self and others' Neuroticism judgments and the 17 neuroticism-relevant RBQ items. The resulting values are listed in Table 4. For Neuroticism, the self (mean = .03) and others' judgments (means ranged from .01 to .04) were generally not correlated with the behaviors examined here.

We also created composite extraversion behavior and neuroticism behavior scores, using methods consistent with those employed to create the emotion composites. After reversing items where theoretically appropriate, participants' RBQ ratings were first transformed to z-scores and then summed. The alpha reliability for the extraversion behavior composite was .67 and the reliability for the neuroticism behavior composite was .82. These scores were then correlated with self and others' personality judgments and the resulting coefficients were compared using t-tests for dependent correlations. Again, the overall pattern of results remained the same (see the bottom row of Tables 3 and 4). The correlations for the self-judgments were consistently higher in magnitude than the correlations for others' for extraversion-relevant behavior but not for neuroticism-relevant behavior. The correlation between the self-judgment of Extraversion and the extraversion behavior composite was .36 and the correlations for others ranged between .13 and .15. The corresponding coefficients for Neuroticism judgments and the neuroticism behavior composite were .07 for self-judgments but ranged between .02 and .04 for others' judgments. Therefore, when self-judgments were compared with the judgments of others to predict behavior, targets were more accurate than others for some, although not all, of the behaviors examined in this study.

Table 4
NEO Neuroticism Correlates of Riverside Behavioral Q-Sort (RBQ)
Items for Self, Individual Informants, and Informant Composite
(*N* = 82)

RBQ Item	Self	Informant #1	Informant #2	Informant Composite
Appears to be relaxed and comfortable (R)	−.10	−.15	−.05	−.12
Exhibits social skills (R)	−.08	.04	.00	.03
Compares self to others	.01	−.14	−.17	−.19
Expresses insecurity	−.06	−.20	.05	−.08
Shows physical signs of tension/anxiety	.09	−.01	.05	.03
Exhibits a high degree of intelligence (R)	−.01	−.31*	.03	−.16
Seeks reassurance from partner	.01	.04	.00	.02
Acts irritated	.11	−.20	−.03	−.14
Expresses hostility	−.04	−.09	.02	−.04
Behaves in a fearful or timid manner	−.04	−.04	.05	.01
Expresses guilt	−.11	−.02	−.05	−.04
Says or does interesting things (R)	−.01	−.05	−.13	−.11
Blames others	−.01	−.02	.00	−.01
Expresses self-pity/victimization	.04	.13	.04	.10
Behaves in a cheerful manner (R)	−.19	.07	−.04	.01
Gives up when faced with obstacles	.13	.07	.14	.12
Speaks fluently and expresses ideas well (R)	−.08	−.15	−.01	−.09
Average	.03	.04	.02	.01
Neuroticism Behavior Composite	.07	.02	.04	.04

Note. Average and composite informant values shown in boldface are significantly different from the average self value at *p* < .05, two-tailed. All averaging of correlation values employed Fisher's *r* to *z* conversion.
(R): Sign of correlation reversed for averaging.
* *p* < .05.

DISCUSSION

The present results demonstrate that the accuracy of a personality judgment may depend, to some extent, on both the unique perspective of the person who provides the judgment as well as the nature of what is being predicted. When the criterion for accuracy was the prediction of emotional experience in daily life, the target individual had the consistently better perspective. When the criterion for accuracy was behavioral prediction, however, the picture was more mixed.

The Importance of Perspective and Cue Availability

When our findings for emotion and behavior are considered together, they could easily be interpreted as implying that self-judgments are simply more accurate than the judgments of others. This would fit with the field's long-held assumption that the target individual has access to the most information regarding his or her personality and is, consequently, the more accurate judge. We would urge caution, however, before drawing such a conclusion. The differences between our behavior findings and those of Kolar et al. (1996) imply that researchers should not assume that a particular judge will *always* have the better perspective on personality. We believe that the results of the two studies suggest that a judge's perspective, which likely results in differential access to cues relevant to the characteristic to be judged, is an important factor in the judgment process. If one is interested in predicting internal aspects of personality such as emotional experience, the self's perspective does seem consistently better. If one is judging traits that are manifest externally and attempting to predict overt social behaviors, however, then the perspective or either the target individual or others may yield accurate judgments. Although we concur with Kolar et al. (1996) that the difference in perspective is a likely explanation for self-other differences in accuracy, we believe this study suggests there is more to the story. It is likely the perspective and access that each type of judge has for the *particular* aspect of personality being judged that is important. If we are to obtain valid judgments of personality that can predict diverse behaviors and phenomena, then we may need to give careful consideration to the selection of the appropriate judge of personality *for the phenomena under study*.

Front Row Seats, Backstage Passes, and the
View From the Inside

These results clearly show that the target individual does have the best perspective from which to judge personality characteristics relevant to emotion. Others may be given considerable access to relevant cues regarding our emotional experience and may even occupy choice front row seats, positions that provide a particularly advantageous perspective for judging the target's external expressions of emotion. The individual's privileged access to inner emotional experience, however, combined with a larger quantity of information, results in more accurate self-judgments. It seems that for the prediction of emotion, the target individual not only has the best seat in the house, he or she has a backstage pass.

Implications for identifying moderators of accuracy.   These findings raise possibilities for additional research regarding the process of judgment as proposed by the RAM. If the self's consistent advantage for emotion is largely attributable to availability and the different access self and others have to emotional experience, then one would expect that others who are given more access to what is usually a private experience should be more accurate. The results of at least one study suggest that if the self discusses thoughts and feelings with an interaction partner, and consequently increases the availability of cues for others, then the judgments of the interaction partner become more accurate (Andersen, 1984). Self-disclosure of emotion effectively takes others behind the scenes and, consequently, should generally enhance others' judgments of internal aspects of personality.

   If others rely on external expressions of emotion to form trait judgments, then the accuracy with which an individual target is judged also may be dependent upon the emotional expressivity of that target for the relevant emotion. According to RAM, "individual differences in the tendency to be judged accurately are a matter of cue availability and relevance" (Funder, 1995, p. 661). If we can obtain measures of individual differences in emotional expressivity for particular traits in everyday settings, then we may be able to identify yet another individual difference that affects target judgability.

   Emotional expressivity could be expected to moderate judgmental accuracy for groups as well as individuals. For example, we would expect the magnitude of self-other differences in accuracy for prediction of

emotion to change with cultural display rules. In cultures where the expression of negative emotion is not discouraged, others may do better for negative emotion than they did in this study. Where it is more strongly discouraged during socialization, then others will likely do worse.

It is also possible that we would find expressivity to moderate gender differences in target judgability as well. For positive emotions, one would not likely see large gender differences. Others should do quite well for positive emotions for both males and females because the expression of positive emotions is encouraged in both. However, we would expect to see gender differences for specific negative emotions. For men, the judgment of sadness and fear would likely be more difficult for others. Men are socialized in this culture to not show these emotions, even if they feel them. Thus, others' access to these negative emotion cues may be quite limited for male targets. We might expect to see a corresponding gender difference with anger in women for similar reasons.

Front Row Seats and We Still Had Trouble Seeing: The View From the Outside

Kolar et al. (1996) found that others occasionally might occupy an advantageous position from which to observe behavior. Figuratively, they are seated in the front row and would seem to have a clear view of the action on stage. Consequently, we expected to find that others were equally as accurate as and sometimes more accurate than the target individual when the criterion was behavioral prediction. That hypothesis was not supported by our data. Although there were no significant differences between self and others' judgments for neuroticism-relevant behaviors, surprisingly, we found that the self-judgments were more accurate for the prediction of extraversion-relevant behaviors. Additionally, unlike Kolar et al. (1996), we did not find that aggregating the judgments of multiple others eliminated the self-other differences in accuracy. In the present study, the composite judgment of two informants was not more accurate than a single self-judgment. The presumption that the aggregate judgment of two others, which has the opportunity to be more reliable, would be more accurate than a single judgment was not supported by our results. In fact, the addition of a second informant did not appreciably improve the size of the correlations for others. Despite the likely improvement in reliability for the informant composite, no substantial improvement in the validity or accuracy of others' judgments

was noted for the specific behaviors examined here. When these results are considered along with those of Kolar et al. (1996), it becomes clear that the question of who has the better perspective on personality for the prediction of behavior, the individual or others, remains to be answered definitively.

The lack of complete replication of the Kolar et al. behavior findings could be due to a number of factors. We examined a much smaller set of behaviors and personality characteristics, choosing to focus our attention on those specific behaviors that were theoretically relevant to our selected personality factors of Extraversion and Neuroticism. The RBQ, although designed to capture a diverse set of social behaviors, was not designed to be a comprehensive list of all personality-relevant behaviors. Restricting our attention to the RBQ cognates of the California Q-set items factor analyzed by McCrae, Costa, and Busch (1986), while a reasonable way to proceed, may have inadvertently omitted other relevant behaviors. One could argue that if we had chosen to examine a broader list of behaviors, as was done in the Kolar et al. study, we might have found that the advantage for targets' self-judgments dissipates. Examination of the full data set, however, essentially ruled out this possibility.[8] The advantage for self-judgments held across the entire list of RBQ items.

The advantage of self-judgments for the prediction of extraversion-relevant behaviors raises some interesting questions that will require additional investigation. Researchers have frequently found that others' judgments become more accurate, or at least tend to agree more, for "good traits" (Funder, 1995), or personality traits with available and relevant behavioral indicators (Borkenau & Liebler, 1992; Funder & Dobroth, 1987; Levesque & Kenny, 1993). Thus, one would logically expect that others' judgments would be equally as accurate as if not more accurate than the self-judgments for extraversion. Because extraversion possesses numerous available, highly visible indicators (e.g., "is talkative," "initiates humor"), it tends to be a characteristic that is easier for others to judge whereas neuroticism, with fewer available and generally less visible indicators, tends to be more difficult. Although an informal perusal of Tables 3 and 4 shows that others' judgments resulted in slightly higher item correlations for extraversion than for neuroticism, as one would expect for a trait with more visible indicators, they remained less accurate on average than the self-judgments. A recent study by Gosling,

8. Available at http://www.psych.ucr.edu/faculty/funder/RAP/perspweb.htm.

John, Craik, and Robins (1998), however, found that observers did not have higher interjudge agreement for extraversion-relevant behaviors (acts). Our study employed both others' trait judgments of extraversion, which tend to yield greater other-other agreement for extraversion, as well as others' judgments of extraversion-relevant behaviors, which may not. Clearly additional research is needed that examines why higher other-other agreement for the trait of extraversion does not necessarily result in higher agreement for extraversion-related behaviors and, ultimately, in greater accuracy.

It is also possible, and perhaps more likely, that another "good trait" moderator of accuracy, relevance, accounts for the current findings (Funder, 1995). It is possible that the list of behaviors we examined included behaviors that, if observed by informants in everyday settings, are behavioral cues that tend to be ambiguous because they may be relevant for more than one trait. These cues are not ambiguous because different people define or interpret them in different ways, rather they are ambiguous because they are relevant to more than one trait; they are *valid* indicators of multiple traits. For example, if Jason "exhibits social skills," does that mean he or she is extraverted or emotionally stable? If Mary "gives up when faced with frustration," is she neurotic or introverted? If John "expresses warmth," "seems likable," and doesn't "keep his partner at a distance," is he agreeable or extraverted? Such ambiguity regarding the relevant cues could easily affect the judgment of even a highly visible trait. A perusal of the factor analysis results from McCrae, Costa, and Busch (1986) suggests that the extraversion-relevant behaviors used here, although definitely related to extraversion, may not be relevant *only* to extraversion. Many of the items also loaded on neuroticism and agreeableness in McCrae et al.'s analysis. Both neuroticism and agreeableness tend to be traits that are difficult for others to judge because social desirability, availability, and cue ambiguity restrict the ability of judges to detect and utilize relevant cues (Funder & Dobroth, 1987; Gosling et al., 1998; Hayes & Dunning, 1997). Although we cannot claim that this is the definitive explanation for the self-other difference obtained for extraversion-relevant behavior in this study, it is an intriguing possibility and one that warrants additional investigation.

Catching Glimpses of Personality

Over 60 years ago, Gordon Allport described the limits of our under-standing of others, stating that

> [n]o person can understand any other person completely because no human being shares directly the motives, thoughts, and feelings of another. The only self to which we have immediate access is our own. Knowledge of other people comes to us indirectly and in fragments. At best we catch glimpses of one another. Yet we try ardently to bridge the chasm between mind and mind, for our happiness and survival depend on correct judgments of persons. (Allport, 1937, p. 497)

Researchers have been wise to heed Allport's general conclusion regarding our limitations when judging the personality of another. The self is often the person with the best access to information relevant to personality judgment. Research, however, has sometimes seemed to miss an important idea: We may catch glimpses of one another that could allow us to make correct judgments about others. The results of this study indicate that the informational value of these glimpses may be importantly related to both our perspective and the nature of what we are trying to see.

REFERENCES

Allport, G. W. (1937). *Pattern and growth in personality.* New York: Holt, Rinehart, and Winston, Inc.

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.

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.

Block, J. (1978). *The California Q-set (Form III).* Palo Alto, CA: Consulting Psychologists Press, Inc.

Bolger, N., & Schilling, E. A. (1991). Personality and the problems of everyday life: The role of neuroticism in exposure and reactivity to daily stressors. *Journal of Personality,* **59**(3), 355–386.

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

Bradley, B. P., Mogg, K., Perrett, A., & Galbraith, M. (1993). The effect of depressed mood on personality measures. *Personality and Individual Differences,* **14**(4), 599–601.

Buck, R. (1984). *The communication of emotion*. New York: The Guilford Press.

Cacioppo, J. T., Klein, D. J., Berntson, G. G., & Hatfield, E. (1993). The psychophysiology of emotion. In M. Lewis & J. M. Haviland (Eds.), *The handbook of emotion* (pp. 119–142). New York: The Guilford Press.

Cantor, N., Norem, J., Langston, C., Zirkel, S., Fleeson, W., & Cook-Flannagan, C. (1991). Life tasks and daily life experience. *Journal of Personality, 59*(3), 425–451.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Costa, P. T., Jr., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology, 38,* 668–678.

Costa, P. T., Jr., & McCrae, R. R. (1984). Personality as a life-long determinant of well-being. In C. Malatesta & C. Izard (Eds.), *Emotion in adult development* (pp. 141–157). Beverly Hills, CA: Sage.

Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.

Craik, K. H. (1986). Personality research methods: An historical perspective. *Journal of Personality, 54*(1), 18–51.

Creed, A. T., & Funder, D. C. (1998a). The two faces of private self-consciousness: Self-report, peer-report, and behavioral correlates. *European Journal of Personality, 12*(6), 411–431.

Creed, A. T., & Funder, D. C. (1998b). Social anxiety: From the inside and outside. *Personality and Individual Differences, 25*(1), 19–33.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience sampling method. *Journal of Nervous and Mental Diseases, 175,* 526–536.

Diener, E., & Larsen, R. J. (1993). The experience of emotional well-being. In M. Lewis & J. M. Haviland (Eds.), *The handbook of emotions* (pp. 405–415). New York: The Guilford Press.

Dutton, D. G., & Aron, A. P. (1974). Some evidence for heightened sexual attraction under conditions of high anxiety. *Journal of Personality and Social Psychology, 30,* 510–517.

Ekman, P. (1973). *Darwin and facial expression: A century of research in review*. New York: Academic Press.

Ekman, P. (1992). Are there basic emotions? *Psychological Review, 99,* 550–553.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face*. Englewood Cliffs, NJ: Prentice-Hall.

Ekman, P., Friesen, W. V., & Ancoli, S. (1980). Facial signs of emotional experience. *Journal of Personality and Social Psychology, 39,* 1125–1134.

Ekman, P., Friesen, W. V., O'Sullivan, M., & Scherer, K. (1980). Relative importance of face, body, and speech in judgments of personality and affect. *Journal of Personality and Social Psychology, 38,* 270–277.

Ekman, P., Roper, G., & Hager, J. C. (1980). Deliberate facial movement. *Child Development, 51*(3), 886–891.

Emmons, R. A., & Diener, E. (1985). Personality correlates of subjective well-being. *Personality and Social Psychology Bulletin, 11,* 89–97.

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. (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**(4), 652–670.

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

Funder, D. C., Furr, R. M., & Colvin, C. R. (1999). *The Riverside Behavioral Q-sort.* Manuscript submitted for publication.

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., & West, S. (1993). Consensus, self-other agreement, and accuracy in personality judgment: An introduction. *Journal of Personality,* **61**(4), 457–476.

Furr, R. M., & Funder, D. C. (1998). A multimodal analysis of personal negativity. *Journal of Personality and Social Psychology,* **74,** 1580–1591.

Gosling, S. D., John, O. P., Craik, K. H., & Robins, R. W. (1998). Do people know how they behave? Self-reported act frequencies compared with on-line codings by observers. *Journal of Personality and Social Psychology,* **74,** 1337–1349.

Gross, J. J., & John, O. P. (1995). Facets of emotional expressivity: Three self-report factors and their correlates. *Personality and Individual Differences,* **19**(4), 555–568.

Hayes, A. F., & Dunning, D. (1997). Construal processes and trait ambiguity: Implications for self-peer agreement in personality judgment. *Journal of Personality and Social Psychology,* **72,** 664–677.

Hjelle, L. A., & Bernard, M. (1994). Private self-consciousness and the retest reliability of self-reports. *Journal of Research in Personality,* **28**(1), 52–67.

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

Hormuth, S. E. (1986). The sampling of experiences in situ. *Journal of Personality,* **54**(1), 262–293.

Izard, C. E. (1991). *The psychology of emotion.* New York: Plenum Press.

Izard, C. E., Libero, D. Z., Putnam, P., Haynes, O. M. (1993). Stability of emotion experiences and their relations to traits of personality. *Journal of Personality and Social Psychology,* **64,** 847–860.

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

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

Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality,* **64**(2), 311–337.

Larsen, R. J., & Ketelaar, T. (1989). Extraversion, neuroticism and susceptibility to positive and negative mood induction procedures. *Personality and Individual Differences,* **10**(12), 1221–1228.

Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology,* **61,** 132–140.

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.

Lewis, M., & Haviland, J. M. (Eds.). (1993). *The handbook of emotions*. New York: The Guilford Press.

Magai, C., & Hunziker, J. (1993). Tolstoy and the riddle of developmental transformation: A lifespan analysis of the role of emotions in personality development. In M. Lewis & J. M. Haviland (Eds.), *The handbook of emotion* (pp. 435–446). New York: The Guilford Press.

McCrae, R. R., & Costa, P. T., Jr. (1985). Updating Norman's "adequate taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology,* **49,** 710–721.

McCrae, R. R., & Costa, P. T., Jr., (1991). Adding Liebe und Arbeit: The full five-factor model and well-being. *Personality and Social Psychology Bulletin,* **17,** 227–232.

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,* **54**(2), 430–446.

McLennan, J., Buchanan, J. I., & Bates, G. W. (1994). Neuroticism and negative affect measures as predictors of psychological distress. *Psychological Reports,* **75**(1), 305–306.

Meyer, G. J., & Shack, J. R. (1989). The structural convergence of mood and personality: Evidence for old and new directions. *Journal of Personality and Social Psychology,* **57,** 691–706.

Moskowitz, D. S. (1986). Comparison of self-reports, reports by knowledgeable informants, and behavioral observation data. *Journal of Personality,* **54**(1), 294–317.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology,* **66,** 574–583.

Ormel, J., & Wohlfarth, T. (1991). How neuroticism, long-term difficulties, and life situation change influence psychological distress: A longitudinal model. *Journal of Social and Personality Psychology,* **60,** 744–755.

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.

Ruiz-Caballero, J. A., & Bermudez, J. (1995). Neuroticism, mood, and retrieval of negative personal memories. *Journal of General Psychology,* **122**(1), 29–35.

Saarni, C. (1993). Socialization of emotion. In M. Lewis & J. M. Haviland (Eds.), *The handbook of emotion* (pp. 435–446). New York: The Guilford Press.

Schacter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of the emotional state. *Psychological Review, 69,* 379–399.

Sneed, C. D., McCrae, R. R., & Funder, D. C. (1998). Lay conceptions of the five-factor model and its indicators. *Personality and Social Psychology Bulletin, 24*(2), 115–126.

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

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.

Valins, S. (1966). Cognitive effects of false heart-rate feedback. *Journal of Personality and Social Psychology, 4,* 400–408.

Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin, 96,* 465–490.

Watson, D., & Clark, L. A. (1992). On traits and temperament: General and specific factors of emotional experience and their relations to the five-factor model. *Journal of Personality, 60*(2), 441–476.

Watson, D., Clark, L. A., McIntyre, C. W., & Hamaker, S. (1992). Affect, personality, and social activity. *Journal of Personality and Social Psychology, 63,* 1011–1025.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*(2), 219–235.

Wheeler, L., & Reis, H. T. (1991). Self-recording of everyday life events: Origins, types, and uses. *Journal of Personality, 59*(3), 339–354.

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