

## Accuracy and Bias in Self-Perception: Individual Differences in Self-Enhancement and the Role of Narcissism

Oliver P. John and Richard W. Robins

Accuracy and bias in self-perceptions of performance were studied in a managerial group-discussion task. Ss ranked their own performance and were ranked by the 5 other group members and by 11 assessment staff members. Although the self-perceptions showed convergent validity with the staff criterion, Ss were less accurate when judging themselves than when judging their peers. On average, Ss evaluated their performance slightly more positively than their performance was evaluated by either the peers or the staff; however, this general self-enhancement effect was dwarfed by substantial individual differences, which ranged from self-enhancement to self-diminishment bias and were strongly related to four measures of narcissism. Discussion focuses on issues in assessing the accuracy of self-perceptions and the implications of the findings for individual differences in self-perception bias and the role of narcissism.

Self-insight, or the accuracy of self-perception, has been an issue of long-standing concern to philosophers and social scientists. Among contemporary psychologists, two different points of view predominate. According to one view, perceptions of self are based on a socially shared reality, ensue from the same processes as the perceptions of others, and are best thought of as accurate reflections of behavior and experience. According to the other view, self-perceptions are fundamentally distorted, self-serving, and consistently more positive than is justified by the perceptions of others. In this article, we argue that each of these views is incomplete. In particular, we demonstrate that self-evaluations in a specific situation contain both valid and

biased components and that the nature of the bias varies as a function of individual differences in narcissism: Some individuals show self-enhancement bias, some show self-diminishment bias, and others are relatively unbiased.

For many years, the study of accuracy in self-perception has been impeded by the "criterion problem": the lack of objective criteria against which self-perceptions can be compared. For example, in studies of personality trait ratings, low correlations between self- and peer judgments cannot be attributed conclusively to faulty self-perceptions; in fact, the self-judgments could be more accurate because they may include relevant and valid information not available to the peers (see Funder, 1989). Similarly, some studies showing that self-ratings are more positive than ratings by uninvolved observers have been interpreted as evidence of a harshness bias on the part of observers rather than an enhancement bias on the part of the self (Coyne & Gotlib, 1983).

To examine these issues in a more controlled setting, we measured the accuracy of self-perceptions of performance against a criterion based on assessments by a staff of psychologists. Our experimental paradigm was a standard simulation task often used in managerial assessment programs (Howard & Bray, 1988; Thorton & Byham, 1982); in this task, subjects were assigned to a decision-making group in which they presented, debated, and then reached consensus about the relative merits of six employees nominated for a merit bonus. This design ensured that the same information was available to both the subjects and a staff of trained psychologists who observed the group discussion, and it permitted us to compare three types of performance evaluations: self, involved peers (i.e., individuals also participating in the task), and uninvolved staff psychologists. Using these three sources of data, we tested the following hypotheses: (a) Are evaluations of self less accurate than evaluations of others? (b) Can the lower accuracy of self-evaluations be explained by a general positivity bias unique to the self, as implied by self-enhancement theory? (c) Or, alternatively, are there systematic individual differences in self-perception bias, as suggested by theoretical accounts of narcissism?

---

This article is based in part on an honors thesis submitted by Richard W. Robins to the University of California, Berkeley.

This research was supported, in part, by National Institute of Health Grant MH-49255 and by Biomedical Research Grants 87-20 and 88-24 (University of California) to Oliver P. John. Richard W. Robins was supported by a National Science Foundation Graduate Fellowship.

We are indebted to Kenneth Craik, who initiated the Institute for Personality and Social Research–Master of Business Administration (IPSR-MBA) Assessment Program while he was director of IPSR, and to the other members of the MBA Assessment planning team. We particularly wish to acknowledge John Kamp, who codirected the 1986 and 1987 assessments, and Robert Raskin, who devised the definition of the observer-based narcissism rating and a short version of the Narcissistic Personality Inventory for the MBA Assessment Program. We are also indebted to Stephen G. West and Leona Aiken for their statistical advice and to Jonathon D. Brown, William F. Chaplin, C. Randall Colvin, Kenneth H. Craik, Robyn M. Dawes, David C. Funder, Lewis R. Goldberg, Sarah E. Hampson, Ravenna Helson, Robert Hogan, Eva Klohnen, Mike Kernis, Adam Kremen, Joachim Krueger, Bernadette Park, Delroy L. Paulhus, Myron Rothbart, Gerard Saucier, Stephen G. West, Drew Westen, Kristina Whitney, and Paul Wink for their comments on a draft of this article.

Correspondence concerning this article should be addressed to Oliver P. John and Richard W. Robins, Department of Psychology, University of California, Berkeley 94720.

## Two Views of Accuracy in Self-Perception

The view of self-perception as based in socially shared reality emphasizes the convergence between judgments by self and judgments by others (e.g., Funder & Colvin, 1988). This convergence has been interpreted as evidence for the accuracy of self-reports (e.g., McCrae, 1982). The other major view, in contrast, emphasizes self-enhancing distortions in people's self-perceptions. For example, Taylor and Brown (1988) characterized self-perceptions as exhibiting pervasive and enduring distortions. We first discuss these two views of accuracy in self-perception and then review previous research on accuracy and bias, emphasizing particularly the need for conceptually relevant and psychometrically sound criterion measures against which to compare self-perceptions.

### *Correspondence View*

Some psychologists have argued that self-perceptions derive from essentially the same processes as do perceptions of others (e.g., Bem, 1972; Cooley 1902; Lewis & Brooks-Gunn, 1979; Mead, 1934). Mead (1934) was one of the first to emphasize the social origin of the self-concept: "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" (Mead, 1934, p. 138). According to Bem's (1972) self-perception theory, individuals acquire self-knowledge by observing their own behavior in much the same way as would an observer, particularly when "internal cues are weak, ambiguous, or uninterpretable" (p. 2). If self-perceptions indeed proceed through the same basic processes as perceptions of others, then the way we perceive ourselves should correspond closely with the way we are perceived by others.

This correspondence assumption underlies much contemporary research on the validity of personality trait ratings. The principle of "consensual validation" (see McCrae, 1982) holds that personality self-ratings are valid if they correlate with ratings made by knowledgeable others; the validity of self-reports is thus defined as differential accuracy, that is, subjects' ability to correctly gauge their standing on an attribute relative to others. The consensual validity of self-reports on a wide range of personality traits has been documented in numerous studies, and self-other correlations have often been substantial (e.g., Funder, 1980; Kane & Lawler, 1978; McCrae, 1982; for a summary, see McCrae & Costa, 1989).

However, additional and potentially nonveridical factors can also influence self-perceptions (e.g., Greenwald, 1980; Paulhus, 1990), and therefore the account of self-perception offered by the correspondence view may be incomplete. In a recent study, (John & Robins, 1993), we compared interjudge agreement between two peers with agreement between the self and a peer on a large set of personality traits. Although self and peers showed agreement for almost all the traits, as predicted by the correspondence view, we also found an important difference: Self-peer agreement was lower than peer-peer agreement for traits that are evaluatively extreme (i.e., ego involving). We speculated that this difference may be due to individual differences in narcissistic self-enhancement. That is, when judging themselves

on evaluative traits, narcissistic individuals will experience a threat to their self-worth and bolster their self-image by perceiving themselves more positively than they are seen by others, whereas this should not be true for relatively modest, nonnarcissistic individuals.

These findings, along with those of others, suggest that the self as a judge differs from judges who are peers or uninvolved observers, especially when the attribute is ego involving (see Greenwald, 1982; Miller, 1976; Tesser & Campbell, 1982). More generally, these findings illustrate the importance of affective factors in judgments about the self and raise the possibility that such factors may decrease the accuracy of self-perceptions.

### *Distortion View*

Most self-concept theorists assume that people are motivated to maintain and enhance their self-esteem (e.g., Allport, 1937; Greenwald, 1980; James, 1890; Rogers, 1959; Tesser, 1988). Recently, Taylor and Brown (1988) have interpreted a wide range of research findings as evidence of pervasive, enduring, and systematic departures of self-conceptions from reality, presumably stemming from the basic motive toward self-enhancement.<sup>1</sup> Taylor and Brown's (1988) conclusions are based on evidence from research in three different domains: unrealistically positive views of the self, illusions of control, and unrealistic optimism. The present research concerns only the first domain—unrealistically positive self-perceptions.

The argument that self-perceptions are unrealistic requires a criterion—a measure of reality—against which self-perceptions can be compared. Given the absence of a single objective standard for complex social behaviors, social scientists have typically assumed that the social consensus (i.e., the aggregated perceptions of others) provides an acceptable criterion for social reality. Surprisingly, only a few studies have directly compared an individual's self-perceptions with a social consensus criterion (i.e., perceptions of that individual by others). One of these studies (Lewinsohn, Mischel, Chaplin, & Barton, 1980) is frequently cited as a demonstration of illusory self-enhancement. However, this study has been criticized for various methodological reasons (Coyne & Gotlib, 1983; Gotlib & Meltzer, 1987), and therefore, we examine it in some detail.

### Past Studies of Self-Enhancement Bias Relative to Ratings by Others

Lewinsohn et al. (1980) compared personality judgments by self and three "objective observers" (p. 210), using three groups

<sup>1</sup> In studies of bias in self-perception, accuracy is typically defined in terms of directional discrepancies between self-perceptions and an external criterion; for example, self-perceptions are considered to be self-enhanced if they are more positive than the criterion. In contrast, psychologists concerned with the validity of self-perceptions tend to focus on differential accuracy (i.e., correlations). These two measures of accuracy are conceptually distinct and statistically independent; that is, self-perceptions could show considerable mean-level discrepancies from a criterion but still be highly correlated with that criterion (see Cronbach, 1955).

of subjects (clinically depressed, nondepressed psychiatric, and normal). Subjects interacted informally with four or five other participants in an unstructured group session and were observed through one-way mirrors by three undergraduates serving as observers. On the basis of their behavior in the group discussion, subjects rated themselves and were rated by the observers on 17 desirable personality traits; analyses of the mean of these 17 ratings showed that the self-ratings were significantly more desirable than the average rating by the three observers. Moreover, an interaction effect indicated that the self-ratings of the depressed group were more similar to the observer ratings than the self-ratings of the nondepressed subjects, a finding that has been interpreted as "depressive realism."

Interestingly, however, the magnitude of the interaction effect, which accounted for 2% of the variance, was dwarfed by a large main effect of rater (i.e., self vs. observer), which accounted for 17% of the variance. In particular, the observer ratings were more negative than the self-ratings of all three groups of subjects, even depressed subjects. Coyne and Gotlib (1983) therefore suggested an alternative interpretation of these self-observer discrepancies: What Lewinsohn et al. (1980, p. 210) described as a "halo or glow that involves an illusory self-enhancement" could also be a harshness bias on the part of the observers. Observer harshness is reflected in the overall elevation (or level) of a judge's ratings across subjects, a potential confound in the assessment of accuracy about which Cronbach (1955) warned the field nearly 40 years ago. Thus, the observers may be the source of the bias rather than the self. In fact, the correlation between the mean observer ratings and mean rankings by the five peers was only .40; the size of this validity correlation seems low for an "objectively" measured criterion variable.<sup>2</sup> In conclusion, the observers in the Lewinsohn et al. study may have provided an unreliable and possibly biased measure of reality.

Gotlib and Meltzer (1987) provided empirical support for Coyne and Gotlib's (1983) observer-harshness interpretation of the Lewinsohn et al. (1980) study. Mildly depressed and nondepressed subjects participated in an unstructured dyadic interaction and then rated themselves and the peer with whom they had interacted on the 17 desirable traits used in the Lewinsohn et al. (1980) study; videotapes of the interaction were later rated by a single undergraduate observer. Subjects' self-ratings were not enhanced relative to the peers' ratings; in fact, the peers' ratings were slightly more positive than the self-ratings. Moreover, the observer's ratings were generally more negative than either self or peer ratings, for both nondepressed and depressed subjects.

J. D. Campbell and Fehr (1990) reported two studies that provide the best-documented failure to find a self-enhancement bias relative to ratings by other subjects participating in the same interaction. In their first study, high and low self-esteem subjects interacted in same-sex pairs for 15 min and then rated their own and their peer's (i.e., their partner's) behavior on 10 desirable and 10 undesirable characteristics. When their ratings were compared with those of their peers, neither high nor low self-esteem subjects showed a self-enhancement bias; in fact, the low self-esteem subjects showed a self-diminishment bias, rating themselves more negatively than they were rated by their

peers. In the second study, the interactions were also observed through a one-way mirror by two undergraduate observers. For the peer ratings, the findings were the same as in the first study; the two uninvolved observers, however, rated both high and low self-esteem subjects more negatively than the subjects rated themselves. Overall, the findings suggest that uninvolved observers make harsher ratings than do either self or peers. J. D. Campbell and Fehr (1990) concluded that the evidence for self-enhancement bias depends entirely on "whose impressions" (p. 128) are used as the criterion for accuracy.

In summary, the available evidence is inconclusive and does not support the general claim that "most individuals see themselves as better than others see them" (Taylor & Brown, 1988, pp. 195–196): When self-ratings were compared with ratings by others involved in the same task, there was no evidence of self-enhancement bias. Moreover, research showing discrepancies between self-ratings and ratings by uninvolved observers may be interpreted as self-enhancement bias, observer harshness, or both. Note that in all of these studies subjects were given Likert rating scales, which allow for individual differences in scale usage; rankings, in contrast, would control for such differences in scale usage.

More generally, the studies reviewed here serve to illustrate the importance of the criterion problem in research on accuracy and bias. One can hardly speak of "objective observers" when one, two, or three undergraduate students are asked to make judgments about global personality traits on the basis of a brief interaction. In these circumstances, it is unclear who is correct and who is biased—the self or the observers? Even when judges are asked to rate overt behavior during a specific task, researchers need to demonstrate that observer judgments provide a reliable and valid criterion measure against which the self-judgments can be compared. We therefore chose a task in which observers can reliably judge subjects' behavior and used a large number of psychologists who were trained to assess performance in this task.

### Self-Enhancement Bias: General Law and Individual Differences

Some psychologists discuss illusory self-enhancement as if it were a general law of human behavior applicable to all normal, psychologically healthy individuals: Taylor (1989, p. 7) concluded that "normal human thought is marked not by accuracy but by positive self-enhancing illusions"; Paulhus and Reid (1991, p. 307) suggested that "the healthy person is prone to self-deceptive positivity"; and Greenwald and Pratkanis (1984, p. 139) asserted that self-enhancing biases pervade the "self-knowledge of the average normal adult of (at least) North American culture," with depressed individuals representing the sole exception.

According to Kurt Lewin (1935), a general law implies that there is a universal intraindividual process (e.g., a motivation to

<sup>2</sup> Lewinsohn, Mischel, Chaplin, and Barton (1980) did not test the self-enhancement effect with the peer criterion because they had obtained peer rankings rather than ratings.

enhance self-esteem) that causes the observed phenomena (e.g., self-enhancement bias) described by the law. What type of evidence supports a general law? Aggregate data analyses, although informative, are not sufficient because intraindividual processes cannot be inferred from sample-level statistics. Yet, most studies do not report the percentage of subjects who actually showed self-enhancement bias and the percentage who did not. Is it 5%, 10%, or 30% of "normal" nondepressed subjects who show no self-enhancement or even self-diminishment? Lewin (1935) once remarked that "It is no longer possible to take exceptions lightly" (p. 24). In the present article, therefore, we report the percentage of subjects who show self-enhancement bias, those who show self-diminishment bias, and those who show neither.

Many researchers do not share the view that self-enhancement bias is a general law of human behavior. Numerous studies have examined the role of individual differences in self-esteem on self-enhancement processes. Compared with low self-esteem individuals, high self-esteem individuals are more likely to make self-serving attributions (e.g., Miller & Ross, 1975), more likely to engage in compensatory self-enhancement after receiving negative feedback (e.g., Baumeister, 1982), more likely to underestimate consensus for their perceived abilities and overestimate consensus for their perceived deficiencies (e.g., J. D. Campbell, 1986), and more likely to derogate sources of negative feedback (e.g., Baumgardner, Kaufman, & Levy, 1989). For example, in a study of reactions to feedback and compensatory self-inflation, Baumgardner et al. (1989) found that high self-esteem subjects show more self-enhancement than low self-esteem subjects in their response to bogus personality feedback; specifically, high self-esteem subjects were more likely to derogate sources of negative feedback and compliment sources of positive feedback. In all, these findings demonstrate that "some individuals self-enhance more than others" (Baumgardner et al., 1989, p. 907).

Note, however, that these studies focus on self-enhancement processes rather than on the accuracy of people's self-perceptions. In contrast to the rich literature on self-enhancement processes, few studies have examined individual differences in self-enhancement bias, measured against an explicit accuracy criterion such as judgments by others. Some previous research suggests that individuals who have low self-esteem (J. D. Campbell & Fehr, 1990) or are depressed (Noles, Cash, & Winstead, 1985) may not only self-enhance less but may actually see themselves more negatively than they are seen by others. The present research builds on and adds to this research by examining another individual-difference variable, namely narcissism.

Although the relevance of narcissism for an individual-differences account of self-enhancement bias seems rather obvious, the construct has not yet been examined in studies of self-perception accuracy against observer criteria. All clinical accounts of narcissism (e.g., Freud, 1914/1953; Kernberg, 1975; Kohut, 1971; Millon, 1981) concur that narcissistic individuals hold unrealistically exaggerated beliefs about their abilities and achievements. According to the criteria specified in the *Diagnostic and Statistical Manual of Mental Disorders*, third edition, revised (*DSM-III-R*, American Psychiatric Association, 1987), the defining characteristics of the narcissistic per-

sonality include "a grandiose sense of self-importance" and a tendency to "exaggerate their accomplishments and talents, and expect to be noticed as 'special' even without appropriate achievement" (pp. 349–350). Both theoretical accounts of narcissism and the *DSM-III-R* definition lead us to predict that narcissistic individuals will show a general tendency toward self-enhancement bias, particularly when the context is evaluative and ego involving and when failure would be threatening (e.g., Emmons, 1987; Raskin, Novacek, & Hogan, 1991a, 1991b). Individuals with intermediate and low levels of narcissism, in contrast, should not show self-enhancement bias; instead, they may be accurate or even show evidence of self-diminishment bias.

## Present Study

### Design

We examined accuracy and bias in subjects' self-evaluations of their performance in a group discussion task; the performance of each subject was also evaluated by the 5 peers in the group and by a staff of 11 trained psychologists. The study was designed so that the staff assessments would provide an acceptable criterion for self-perception accuracy. We chose the interaction situation (a simulation of a managerial committee meeting) and the judgment dimension (the individual's performance during the group discussion) so that, in principle, all judges would have the same opportunity to observe the relevant behaviors and would have available the same information on which to base their judgments. The definition of the task and the judgment dimension made "privileged" or prior knowledge (e.g., intentions, motives, and past behavior) irrelevant to the judgments.

To permit us to directly compare the self, peer, and staff evaluations, both subjects and assessment staff ranked the six group members in terms of their relative performance. In contrast to ratings (which do not have an absolute, or at least consensually shared, anchor point or norm), the ranking format provides an explicit context of comparison for the self-evaluations (i.e., the five other participants), and it also eliminates any potential differences in scale usage between the assessment staff and the subjects involved in the task. Therefore, differences between self-evaluations and staff assessments cannot be attributed to general observer harshness, because each assessor and each subject had to use each of the six ranks and thus both were constrained to the same overall mean evaluation.

The ranking format also makes it less likely that self and other judges use different scales of comparison; differences in the availability of personal history data might lead the observer to evaluate an individual's behavior relative to others, whereas the self might judge with reference to previous behavior rather than the behavior of others (Jones & Nisbett, 1972, p. 85). In the present study, both subjects and assessment staff had to answer the same question for each group member: "How well did this person perform during the 40-minute discussion compared to the other five members of the group?"

### Hypotheses

We expected to find substantial validity (i.e., differential accuracy) for the peer and self-evaluations because the informa-

tion relevant to the judgment dimension was fairly clear and observable and occurred within a limited time period. We predicted, however, that subjects' evaluations of their own performance would be less accurate than their evaluations of the performance of their peers; in other words, correlations between self-evaluations and staff assessments would be lower than correlations between peer evaluations and staff assessments. Following Taylor and Brown (1988), we expected that subjects would exhibit a self-enhancing bias in their evaluations of their own performance, relative to the evaluations of their performance by the assessment staff and by their peers. We also predicted, however, that individuals would differ substantially and systematically in the extent to which they showed self-enhancement, with some individuals overestimating and others underestimating their performance. Finally, theoretical accounts of narcissism as a dimension of normal personality and recent findings of Raskin et al. (1991a, 1991b) suggest that individual differences in self-enhancement bias would be predictable from both self-report and observer-based measures of narcissism.

## Method

### Subjects

Subjects participated in a combined personality and managerial assessment program conducted at the Institute of Personality and Social Research (IPSR). Subjects were master of business administration (MBA) students at the University of California at Berkeley who had volunteered to participate in a weekend assessment. They ranged in age from 21 to 41 years, with a median age of 27. The average subject had more than 3 years of work experience before entering the MBA program. The 12 subjects who participated in each assessment weekend were randomly assigned to one of two discussion groups, with the only constraint that there were about equal numbers of men and women in each group.<sup>3</sup> The present analyses are based on 102 subjects (56 men and 46 women), who participated in 17 groups of 6 subjects.

### Group Discussion and Performance Evaluations

During the assessment weekend, subjects participated in a group discussion procedure simulating the meeting of a compensation committee in a large company. The six subjects were seated at a round table in randomly predetermined seats. The status of all subjects was equal, and no leader was designated. The subjects received a realistic written summary of the employment backgrounds of six candidates for a merit bonus, including salary, biographical information, and appraisals of prior job performance. Each subject was assigned the role of supervisor of one of the candidates and was instructed to present that candidate's case at the committee meeting. The instructions to the participants stressed that they should try to achieve three goals in this task: (a) consensually allocate a fixed amount of bonus money in a way that was in the best interests of the company, (b) obtain as large a bonus as possible for the candidate they were representing, and (c) complete the committee assignment within the 40-min meeting time. Thus, effective performance required behaviors that promoted the achievement of these three goals.

At the end of the committee meeting, we obtained self, peer, and staff evaluations of each subject's performance in the group discussion. Subjects ranked themselves and the five other participants (peers) on the extent to which each of them had contributed to the overall effectiveness of the group; thus, subjects had to compare their own performance with that of the five other participants. The groups were also observed by at least 11 staff assessors, each of whom independently ranked the subjects

on the same dimension as the self and peer judges. Members of the assessment staff varied from weekend to weekend and included experienced personality, industrial-organizational, and clinical psychologists associated with IPSR, as well as graduate students trained to evaluate performance in this task.

The availability of staff assessments provided us with an ecologically valid criterion against which the MBA students' self-perceptions could be compared. The specific managerial simulation task we used was developed by an organizational consulting firm and is widely used in industry to evaluate managerial performance. Staff assessments of performance in this task have been shown to predict a variety of important work outcomes, including job performance and career advancement (Howard & Bray, 1988; Thorton & Byham, 1982), and thus represent an appropriate accuracy criterion for the present research.

The group discussion is a realistic task for the MBA students, many of whom had already engaged in similar assignments in their jobs. The students generally consider their performance in this task informative of their career potential and thus were highly motivated to perform well. Observations of the subjects during and after the task suggest that they found it ego involving.

### Measures of Narcissism

Narcissism is a complex construct that has been difficult to conceptualize and measure (Raskin & Terry, 1988). Thus, to test the generalizability of our findings, we used four different measures—two obtained from observers (i.e., the assessment staff) and two from self-reports.

*Observer-based measures.* Subjects interacted with each other and with members of the assessment staff for 2½ days. Their activities ranged from informal breakfasts and lunches with staff members to more structured tasks, such as managerial assessment exercises, interviews, and a game of charades. After observing the subjects throughout the weekend, assessment staff members recorded their impressions and observations of the subjects using a variety of personality assessment methods. These data provided us with two observer-based measures of narcissism.

The first was a direct rating of narcissism by all assessment staff members ( $\alpha = .86$ ) using the *DSM-III-R* definition, "self-admiration that is characterized by tendencies toward grandiose ideas, fantasized talents, exhibitionism, and defensiveness in response to criticism; and by interpersonal relations characterized by feelings of entitlement, exploitativeness, and lack of empathy." A second observer-based measure of narcissism was obtained from the California Adult Q-set (CAQ; Block, 1961/1978). The subjects were described using the 100 CAQ items by those five assessment staff members with whom they had the most intensive contact during the assessment weekend (e.g., an in-depth life-history interview). These five Q-sorts were pooled to yield a composite personality description of each subject. To derive a narcissism score from these CAQ descriptions, we used Wink's (1991) narcissism prototype, a composite of the CAQ descriptions of the prototypical narcissist by nine clinical-expert judges. Specifically, we computed a CAQ Narcissism Prototype score for each subject by correlating the composite CAQ description of the subject with Wink's narcissism prototype across the 100 CAQ items (see Block, 1961/1978). These CAQ Narcissism Prototype scores ranged from  $-.50$  to  $.55$  ( $M = -.15$ ,  $SD = .24$ ), with higher scores indicating a more narcissistic personality profile.

*Self-report measures.* Two self-report measures of narcissism were available: a 33-item version of the Narcissistic Personality Inventory (NPI; Raskin & Terry, 1988) and Wink and Gough's (1990) 49-item

<sup>3</sup> We found no significant gender differences in mean levels of self-enhancement bias or in correlations between self-enhancement and narcissism. Therefore, we report analyses only for the combined sample.

Narcissism scale scored from the California Psychological Inventory (CPI; Gough, 1987). The NPI (Raskin & Hall, 1979) was designed for nonclinical populations and is the most widely used and thoroughly researched measure of narcissism. The version of the NPI used in the present sample had a mean of 15.6 ( $SD = 5.3$ ) and a coefficient alpha of .79. The CPI Narcissism scale had a mean of 26.8 ( $SD = 6.6$ ) and an alpha of .79; these values are similar to the college norms reported by Wink and Gough (1990).

The subjects completed the CPI and the NPI several days apart. CPI Narcissism scores and CAQ Narcissism Prototype scores were available for all 102 subjects; NPI scores and *DSM-III*-based observer ratings of narcissism were available for 72 subjects.

## Results and Discussion

First, we examined the reliability and validity of the assessment staff criterion. Second, we tested the hypothesis that self-evaluations would show less convergence with the staff criterion than evaluations of peers. Third, we examined the main effect of self-enhancement bias relative to staff assessments. Finally, we tested the hypothesis that substantial individual differences in self-enhancement bias exist and that they are predictable from narcissism.

### Reliability and Validity of the Staff Assessments

We examined the reliability and validity of the mean staff assessment to justify its use as a criterion for the accuracy of the self-rankings. The 11 staff assessors achieved much better reliability than observers in previous research on accuracy. As shown in Table 1, the staff assessments had near-perfect reliability ( $\alpha = .94$ ). Nonetheless, the staff may have judged performance differently than the subjects participating in the group discussion. To examine this possibility, we correlated the staff assessments with the mean of the five peer evaluations available for each subject. This correlation was .81, indicating substantial agreement between the staff assessments and the subjects' evaluation of their peers. Given that the reliability of the mean peer evaluation was .80, the validity correlation between staff and peers suggests that nearly all of the reliable variance in the peer evaluations was valid. The substantial size of this convergence correlation attests to the validity of both staff and peer evaluations and supports the use of the peer judgments as a second criterion for self-perception accuracy.

Table 1  
*Correlations Among Self, Mean Peer, and Assessment Staff Rankings, Computed Across All 102 Subjects*

Performance evaluation	Assessment staff	Mean peer
Assessment staff	.94 <sup>a</sup>	
Mean peer	.81	.80 <sup>a</sup>
Self	.41	.42

Note. All correlations are significant at  $p < .01$ . Correlations involving the self are set in bold.

<sup>a</sup> Coefficient alpha reliabilities of the composite rankings of the 5 peers and of the 11 staff members.

Table 2  
*Correlations Among Self, Mean Peer, and Assessment Staff Rankings, Computed Across the Six Individuals in Each Group and Then Averaged Across the 17 Groups*

Performance evaluation	Assessment staff	Mean peer
Assessment staff	.95 <sup>a</sup>	
Mean peer	.84	.81 <sup>a</sup>
Self	<b>.58</b>	<b>.64</b>

Note. All correlations are significant at  $p < .01$ . Correlations involving the self are set in bold.

<sup>a</sup> Coefficient alpha reliabilities of the composite rankings of the 5 peers and of the 11 staff members.

### Validity of the Self-Evaluations of Performance

Having established two reliable and valid criteria for accuracy, we used them to evaluate the validity of the self-evaluations. We first correlated the self-rankings with the staff and peer criteria across all 102 subjects, and then we replicated these correlations separately within each of the 17 groups. Across all 102 subjects, both the staff-self and the peer-self correlations exceeded .40 (see Table 1). These findings replicated when we computed convergence correlations across the six subjects within each group and then averaged them across the 17 groups (see Table 2).<sup>4</sup> The size of the correlations in both types of analyses documents the considerable validity (i.e., differential accuracy) of subjects' self-evaluations.

Note, however, that in both Tables 1 and 2 the two correlations involving the self-evaluations were considerably lower than the correlation between the staff and peer evaluations. Although we predicted this differential pattern of correlations, our findings could be explained by the staff and peer criteria both being based on aggregated judgments and therefore being more reliable than the single self-judgment. A strict test of whether subjects' evaluations of themselves were less accurate than their evaluations of their peers requires that we compute pairwise correlations between the self-rankings and the unaggregated rankings by each of the individual staff members and the peer judges.

### Accuracy of Self-Evaluations Relative to Evaluations of Peers

To examine the pairwise correlations among the individual (unaggregated) judges, we computed five self-peer correlations (i.e., one self-judgment with each of the five peers), 11 self-staff correlations, 10 peer-peer correlations, 55 staff-staff correlations, and 55 peer-staff correlations across the 102 subjects. The means of these pairwise agreement correlations, averaged across all possible pairs of judges in each cell, are presented in

<sup>4</sup> The Pearson correlations closely approximated Spearman's rank correlations in all analyses; consequently, only Pearson correlations are reported. In all computations involving correlation coefficients, Fisher's  $r$ -to- $z$  transformation was used.

Table 3  
*Mean Correlations Between Individual Judges (Self, Peer, and Staff Member) Averaged Across All Pairs of Judges in Each Cell*

Judge	Staff member	Peer
Staff member	.60 <sup>a</sup>	
Peer	.49	.45 <sup>a</sup>
Self	.32 <sup>b</sup>	.32 <sup>b</sup>

*Note.* All correlations are significant at  $p < .01$ . Correlations involving the self are set in bold.

<sup>a</sup> Each staff member (or peer) was correlated with each other staff member (or peer); the value in the table indicates the average of these pairwise interjudge correlations. <sup>b</sup> The two correlations involving self-perceptions are significantly lower than the other three correlations ( $p < .001$ ).

Table 3. These correlations suggest two important findings. First, even when neither the self-evaluations nor the evaluations by peers and staff were aggregated, all the agreement correlations were nonetheless significantly different from zero (all  $ps < .001$ ). On average, the single self-ranking correlated above .30 with rankings by individual staff members and individual peers.

Second, the predicted differences among these agreement correlations were also confirmed. In particular, individual-judge agreement between self and peer ( $r = .32$ ) and between self and staff ( $r = .32$ ) was significantly lower than agreement between peer and staff ( $r = .49$ ), as shown by  $Z$  tests for dependent sample correlations ( $p < .001$ ). In variance terms, the convergence between self and peer evaluations accounted for about 10% of the variance, and the convergence between self and individual-staff evaluations also accounted for about 10%. In contrast, peer and staff evaluations had 24% of the variance in common, more than twice the variance each shared with the self-judgments. Apparently, self-evaluations contain more unique variance (i.e., variance that is not shared with the staff assessments) than evaluations of peers. Again, these findings were replicated when we computed the individual-judge correlations across the six subjects within each discussion group and then averaged the correlations across the 17 groups.

Overall, our results indicate that the performance evaluations studied here had considerable reliability and showed substantial convergent validity across self, peers, and assessment staff; even the unaggregated self-evaluations converged to a significant extent with the evaluations of others. Although this finding is consistent with the correspondence view, our second major result is not. In particular, we provided support for our earlier argument (John & Robins, 1993) that perceptions of self are less accurate than perceptions of peers when the dimension judged is ego involving. The present finding, which replicated across all 102 subjects and within the 17 individual groups, is particularly compelling because the same subjects provided both self- and peer judgments. When judging the performance of their fellow participants, our subjects agreed substantially with the assessment staff about the relative standing of the other participants. When judging their own performance, however, the same subjects agreed much less with the assessment staff about their relative standing in the group.

### *Self-Enhancement Bias in Self-Evaluations: Main Effect in the Overall Sample*

To test whether subjects showed self-enhancement in their performance evaluations, we compared their self-rankings with the staff and peer criteria, both across all 102 subjects and separately within each group. On average, the subjects ranked themselves .24 of a rank better than they were ranked by the assessment staff,  $t(102) = 1.67$ ,  $p < .05$  (one-tailed), and .29 of a rank better than they were ranked by their peers,  $t(102) = 2.03$ ,  $p < .05$  (one-tailed). The mean self-ranking was 3.26 ( $SD = 1.28$ ), the mean staff ranking 3.50 ( $SD = 1.35$ ), and the mean peer ranking 3.55 ( $SD = 1.32$ ). To assess the generalizability of the self-enhancement effect across the 17 groups, we computed the difference between the mean of the six self-rankings and the staff and peer criteria separately in each group. In 9 (53%) of the 17 groups, subjects overestimated their performance. In 5 (29%) of the groups, however, they underestimated, and in 3 (18%) groups the mean self-ranking was identical to the mean staff and peer rankings. These findings were identical for the staff and peer criteria.

In general, then, our findings provide some support for the basic premise of self-enhancement theory: On average, subjects overestimated their performance in the group discussion. In contrast to earlier studies, this finding can be attributed unambiguously to self-enhancement because we ruled out the observer harshness explanation by using a ranking (rather than a rating) procedure. Moreover, we obtained essentially the same effect size for the peer evaluations as for the staff assessments. Thus, when the overall sample mean is considered, the present study provides evidence for self-enhancement bias, whether we use staff or peer evaluations as the accuracy criterion.

The size of this general self-enhancement effect, however, was not large (corresponding to about one quarter of a rank or about one fifth of a standard deviation of the rankings). Moreover, the effect held for only about half of the discussion groups, and 47% of the groups showed either perfect accuracy or self-diminishment. These results reveal substantial individual differences that make the overall sample mean a potentially misleading summary statistic.

### *Self-Enhancement Bias in Self-Evaluations: Individual Differences*

If self-enhancement bias is not a general law of self-perception, then we should find a nontrivial number of subjects who fail to show the general self-enhancement effect. These subjects would not necessarily be more accurate in their self-perceptions; some of them might have quite inaccurate self-views, showing unrealistic self-diminishment rather than self-enhancement. In fact, our results showed that both the magnitude and the direction of self-perception bias varied across individuals. For each subject, we computed two individual discrepancy scores—the staff criterion rank minus the self-rank and the peer criterion rank minus the self-rank; these scores had a possible range from +5 to -5. Subjects' scores ranged from +3.0 ranks (overestimation) to -4.7 ranks (underestimation) for the staff criterion and from +4.0 (overestimation) to -4.2 (underestima-

tion) for the peer criterion, indicating that some subjects severely overestimated their performance, whereas others severely underestimated.

The percentage of subjects who over- or underestimated their performance is presented in Table 4 separately for the two criteria. Overall, about 60% of the subjects overestimated, whereas a sizable minority of 38% underestimated. However, many of the subjects were quite accurate in their self-evaluations. As shown in Table 4, more than 50% of the subjects ranked themselves within one rank of the staff and peer criteria; less than 35% overestimated by more than one rank, whereas 15% underestimated by more than one rank.

It is possible that the individual differences observed here are simply random error around the sample mean. To test whether these individual differences are consistent and psychologically meaningful, we related them to four different measures of narcissism.

#### *Self-Enhancement Versus Self-Diminishment Bias: A Function of Narcissism*

The main effect of self-enhancement in a sample is typically indexed by the simple difference between the self and the staff judgment. However, such difference scores have been widely criticized as indexes of individual differences because their reliabilities are low and because they tend to be confounded with the variables that comprise the index (e.g., Cohen & Cohen, 1983; Cronbach & Furby, 1970). In the present study, the simple difference score (keyed in the direction of self-enhancement) would be negatively correlated with the assessment staff criterion, so that individuals receiving poor evaluations by the staff would appear to self-enhance more than individuals receiving favorable evaluations by the staff.

To provide two unconfounded measures of self-enhancement (vs. self-diminishment) bias, we used residual scores computed by regressing either the staff criterion or the peer criterion onto the self-evaluations and retaining the standardized residuals of the self-evaluations. These residual scores represent the variance that remains in the self-evaluations after the variance predictable from the staff or peer criterion has been removed.

Table 4  
*Percentage of Subjects Overestimating and Underestimating Their Performance According to the Assessment Staff and Peer Criteria*

Self vs. other judgments	Criterion	
	Staff	Peer
Overestimate	60%	58%
Same rank	3%	4%
Underestimate	36%	38%
Overestimate > 1 rank	32%	31%
Within one rank	53%	54%
Underestimate > 1 rank	15%	15%

Note.  $N = 102$ .

Table 5  
*Narcissism Predicts Self-Enhancement Bias: Correlations of Observer-Based and Self-Report Measures of Narcissism With Individual Differences in Self-Enhancement Bias*

Narcissism measure	Self-enhancement bias		
	Staff criterion	Peer criterion	Combined criterion
Observer judgments			
<i>DSM-III-R</i> -based rating	.38**	.40**	.40**
CAQ prototype score	.17	.23*	.20*
Observer composite	.34**	.38**	.36**
Self-reports			
Narcissistic Personality Inventory	.33**	.30*	.32**
CPI Narcissism scale	.21*	.23*	.23*
Self-report composite	.34**	.34**	.34**
Total narcissism composite (all four measures combined)	.44**	.48**	.46**

Note. Positive correlations indicate overestimation, and negative correlations indicate underestimation, relative to the staff (or peer) criterion. *DSM-III-R* = *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev., American Psychiatric Association, 1987); CAQ = California Adult Q-set; CPI = California Psychological Inventory. \*  $p < .05$ . \*\*  $p < .01$ .

Given that the staff and peer criteria seem to reflect the "behavioral reality" of subjects' performance (i.e., observable behavioral differences), the residual scores represent the degree and direction of the bias that remains in the self-rankings after the behavioral reality component has been partialled out. Positive values on these residualized indexes reflect relative self-enhancement (i.e., overestimation by the self), and negative values reflect relative self-diminishment (i.e., underestimation by the self).

The following analyses use these two individual-differences measures of self-enhancement versus self-diminishment bias, as well as their composite (i.e., the unit-weighted mean of the staff- and peer-based indices). We tested the narcissism effect in two sets of analyses, computing correlations (a) across all 102 subjects (reflecting individual-differences variance in the entire sample) and (b) separately across the six members in each group (within-group variance), permitting us to replicate the effect in 17 groups. In both analyses, we used the four individual measures of narcissism (the two observer-based measures and the two self-report scales) and three unit-weighted composites: (a) an observer composite combining the *DSM-III-R* rating and the CAQ Prototype score, (b) a self-report composite combining the NPI and the CPI Narcissism scale, and (c) a total composite combining all four measures.<sup>5</sup>

*Analyses across all subjects.* The correlations between self-enhancement and narcissism across all subjects are reported in Table 5. As predicted, the correlations were positive and generally substantial. The narcissism effect held for both the staff- and

<sup>5</sup> All four narcissism measures were significantly intercorrelated. The mean intercorrelation was .44, and the alpha reliability of the composite of all four measures was .74.

Table 6

*Replication of the Correlation Between Narcissism and Self-Enhancement Within the 17 Groups: Means of the Within-Group Correlations and the Percentage of Groups Showing a Positive Correlation*

Narcissism measures	Self-enhancement indices <sup>a</sup>			Groups showing the effect (%)
	Staff criterion	Peer criterion	Combined criterion	
Observer judgments				
<i>DSM-III-R</i> -based rating	.49	.55	.55	100
CAQ prototype score (Wink & Gough, 1990)	.17	.28	.23	70
Observer composite	.36	.46	.43	100
Self-reports				
Narcissistic Personality Inventory (Raskin & Terry, 1988)	.38	.34	.36	83
CPI Narcissism scale (Wink & Gough, 1990)	.19	.22	.21	64
Self-report composite	.33	.36	.34	75
Total narcissism composite	.40	.49	.45	93

*Note.* Values are means of the 17 within-group correlations for the CAQ prototype scores and the CPI Narcissism scale and means of the 12 correlations for the *DSM-III-R*-based narcissism rating, the Narcissistic Personality Inventory, and all composite measures. Composite scores are unit-weighted sums of standard-scored variables. *DSM-III-R* = *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev., American Psychiatric Association, 1987); CAQ = California Adult Q-set; CPI = California Psychological Inventory.

<sup>a</sup> The self-enhancement indices are standardized residuals after removing the mean staff (or peer) variance from the self-ranks. Positive values indicate overestimation, and negative values indicate underestimation, relative to the staff (or peer) criterion.

the peer-based self-enhancement indexes and replicated across both observer and self-report measures of narcissism. The most stringent test of the predicted relation is the correlation between the peer-based self-enhancement index and the observer measure of narcissism because they are based on independent data sources and thus provide a test of the effect that cannot be attributed to method overlap; this correlation was .40 ( $p < .01$ ), demonstrating the substantial generalizability of the narcissism effect across methods. Using the total narcissism composite, the correlations were .48 for the peer-based self-enhancement index, .44 for the staff-based index, and .46 for the combined index (all  $ps < .001$ ).<sup>6</sup>

To examine the nature of the relation between narcissism and self-enhancement bias, we inspected the bivariate scatter plots and found that the relation was approximately linear. In fact, when we divided the sample into thirds on the basis of the total narcissism composite, we found that the group of individuals high in narcissism showed substantial overestimation (.47 in standard score metric), the group of individuals low in narcissism showed substantial underestimation (-.41), and the intermediate group showed neither bias (-.05). Thus, in addition to unrealistic self-enhancement among highly narcissistic individuals, individuals scoring extremely low on narcissism were also unrealistic in their self-perceptions—they were overly modest and showed a self-diminishment bias. Subjects with the most accurate self-perceptions were those in the middle of the narcissism distribution.

*Within-group replication.* The narcissism effect also replicated when we computed correlations within each of the 17 groups and averaged them across the 17 groups. The means of these within-group correlations are presented in Table 6, along

with the percentage of groups showing the narcissism effect. The total narcissism composite and the self-enhancement index correlated .49 for the peer-based index, .40 for the staff-based index, and .45 for the combined index. Moreover, the predicted effect replicated in 11 (93%) of the 12 groups for which the total narcissism composite was available, and 8 of these 12 correlations exceeded .40. Thus, the relation between self-enhancement bias and narcissism is pervasive and strong, whether the data are analyzed across all subjects or separately within each group.

In summary, both types of analyses provided strong support for the hypothesized link between individual differences in self-enhancement and narcissism. This relation generalized across two measures of self-enhancement (vs. self-diminishment) bias, four measures of narcissism, and 93% of the discussion groups. These findings suggest that individual differences in self-enhancement bias are systematic and psychologically meaningful. Moreover, the link with narcissism, a broad and stable personality characteristic, implies that the individual differences in self-enhancement bias measured in the present task reflect, in part, a stable and generalized tendency to see oneself unrealistically positively.

<sup>6</sup> In addition to the self-enhancement bias, it is also informative to consider how well the narcissists performed, contrasting their own self-evaluations with the peer and staff evaluations. Whereas the total narcissism composite was highly related to the self-evaluations of performance ( $r = .44, p < .001$ ), it was not related to either staff or peer evaluations ( $r_s = .12$  and  $.04$ , respectively). In other words, the narcissists thought that they had performed considerably better than the average subject, whereas the observers and peers did not think so.

## Conclusions and Implications

The present research compared self and peer evaluations of performance in an ego-involving task with an accuracy criterion based on staff assessments. Our findings suggest three general conclusions: (a) People are less accurate when evaluating their own performance than when evaluating the performance of others; (b) although many people evaluate themselves more positively than justified, a substantial number evaluate themselves quite realistically, and some even make unrealistically negative self-evaluations; and (c) people whose self-evaluations are the most unrealistically positive tend to be narcissistic. Each of these conclusions is discussed in turn.

### *Issues in Assessing the Accuracy of Self and Other Perceptions*

Consistent with the correspondence view of self and other perception, we found substantial convergence between self-evaluations and the assessment staff criterion. However, we also found that the self-evaluations showed less convergence with the criterion than the peer evaluations. This difference is particularly compelling because the same subjects provided both self- and peer judgments: *When judging the performance of their peers, our subjects agreed substantially with the assessment staff; when judging their own performance, however, the same subjects agreed much less with the assessment staff.*

Although we have demonstrated that the self-evaluations agreed less with the assessment staff, conclusions about accuracy depend on the degree to which the staff assessments can be defended as a plausible measure of reality in the present context. Our case for an accuracy interpretation of our findings is based on four general points. First, as described in the introduction, the task and the dimension to be judged were designed to make observer judgments an appropriate criterion: Prior and privileged information were irrelevant to the judgment, the ranking format constrained the judges to make comparative judgments, and all of the behaviors relevant to making the judgments could be observed by all judges. Second, the ecological validity of staff assessments of performance in managerial simulation tasks has been demonstrated in prior research. Third, the reliability of the staff assessments was near perfect. Fourth, the convergence between the staff assessments and the peer judgments was substantial (approaching the reliability of the peer judgments), thus supporting the validity of the staff criterion.

Is our case for an accuracy interpretation of our findings consistent with theoretical accounts of judgmental accuracy? Three recent conceptions of accuracy concur that observer judgments can provide a useful, albeit imperfect, criterion for assessing the accuracy of social perceptions.

Kenny (1991) noted that "the most valued 'instrument' used by psychologists is the human observer" (p. 156) and provided a definition of accuracy formulated in terms of observer judgments: "the average judgment made by all possible observers of all possible target behaviors" (p. 159). Our operational definition of accuracy closely followed Kenny's theoretical definition: (a) The very high reliability of the staff assessments ( $\alpha = .94$ ) shows that adding even more judges would do little to change

the average judgment defined by Kenny as the accuracy criterion; (b) although we could not study all possible observers, we did use a much larger number of observers than in most previous studies on accuracy and bias, and we included two types of judges (i.e., trained assessors and the participants involved in the task); and (c) all possible target behaviors relevant to the dimension being judged could be observed by all of the judges.

Kruglanski (1989) differentiated three notions of judgmental accuracy: consensus, correspondence, and pragmatic utility. First, our criterion obviously reflects consensus; agreement among the staff members was substantial. Second, the correspondence notion of accuracy refers to the relation between subjects' judgments and a criterion. Kruglanski suggested that it is especially important that subjects and experimenters agree about the appropriateness of the criterion. In the present context, both the MBA students and we (the researchers) agree that staff assessments provide a valid criterion.<sup>7</sup> A related issue is whether assessors and subjects agree about the standards for evaluating effective performance. We expected them to agree because our task is a standard managerial assessment procedure requiring the kind of managerial skills our subjects have learned in their classes and on their jobs. Empirically, we found substantial convergence between the staff assessments and the subjects' evaluations of their peers ( $r = .81$ ), providing some assurance that subjects and assessors understood the task and defined performance in a similar way. Kruglanski's third notion of accuracy involves considerations of pragmatic utility—the adaptive or functional value of the judgment. Previous research shows that assessments of performance in our task by trained assessors predict a variety of important work outcomes, such as job performance and career advancement (e.g., Thornton & Byham, 1982). Thus, the staff assessments provide an ecologically valid criterion that has pragmatic utility for our MBA subjects.

Funder (e.g., 1987, 1989, 1990) has written extensively on accuracy issues, particularly in the context of personality judgments. In general, he argues that "the study of accuracy in judgment is exactly the same thing as measurement validity, where the measurements being validated are interpersonal judgments" (Funder, 1990, p. 208). From this perspective, a personality judgment is accurate to the extent that it agrees with judgments by others and predicts relevant behaviors. Although our subjects made judgments about performance in a specific task rather than about global personality traits, Funder's views are nonetheless pertinent. As we have already noted, the staff assessments show a high degree of consensus, agree with the peer judgments, and, according to previous research, predict future behavior and outcomes in the workplace.

In summary, we believe that our accuracy criterion is both theoretically and empirically well-justified. Nonetheless, some

<sup>7</sup> To verify this claim, we asked 33 MBA students who had participated in an IPSR assessment weekend whether they thought the staff assessments were a valid measure of their effectiveness in the group discussion task. On a 9-point rating scale (1 = *not at all valid*; 9 = *extremely valid*), the mean rating was 7.3 ( $SD = 1.3$ ), and 76% of the subjects rated the staff criterion as at least *quite valid* (i.e., 7 or above). Moreover, subjects thought the staff assessment was more valid than their own self-evaluations,  $t(32) = 3.9, p < .001$ .

psychologists maintain that the subjectivity intrinsic in all social judgments, whether by subjects, peers, or trained assessors, makes them epistemologically questionable and thus unsuitable as a criterion for reality. We agree that our criterion, like all criteria, is only a fallible yardstick for reality. However, we concur with Kruglanski (1993) that the skeptics "may be right in pronouncing objective Truth impossible to attain. But they may risk 'pouring out the baby with the tub water' when they renounce the quest for subjective or intersubjective Truth" (p. 4).

Observer judgments are widely used to assess social behavior in developmental, industrial-organizational, personality, social, and other areas of psychology. All of these studies assume that consensual and construct-validated judgments are accurate, and just like these studies, the present research depends on such judgments. In fact, Taylor and Brown's (1988) argument that self-perceptions are unrealistically positive relies on research using observer and peer judgments as criteria (e.g., Lewinsohn et al., 1980).

### *Evidence for Self-Enhancement Bias*

Do the self-judgments show a general bias toward unrealistically positive self-evaluations? Previous research was open to the alternative interpretation that observer ratings were overly harsh (J. D. Campbell & Fehr, 1990; Coyne & Gotlib, 1983). To eliminate the possibility of observer-harshness effects, we used comparative judgments (i.e., rankings). Nevertheless, the average subject overestimated his or her contribution to the group discussion. Thus, our study provides an unconfounded demonstration of the self-enhancement bias described by Taylor and Brown (1988). However, this bias was relatively small: On average, subjects overestimated their performance by about one quarter of a rank.

Note that this general self-enhancement bias cannot explain the lower accuracy of the self-judgments; if all individuals self-enhanced to the same degree, then the correlation between the self-evaluations and the accuracy criterion would not be affected. Rather, the lower accuracy of judgments about the self must be due to individual differences in self-perception bias.

### *Individual Differences in Self-Enhancement Bias and Narcissism*

In addition to the general tendency toward self-enhancement bias for the sample as a whole, we also found substantial individual differences in both the magnitude and the direction of self-perception biases. In fact, our findings suggest a continuum ranging from extreme levels of self-enhancement on one pole to extreme levels of self-diminishment on the other pole. Only 35% of the subjects showed clear evidence of unrealistically positive self-perceptions, whereas 50% were fairly accurate. The remaining 15% showed evidence of unrealistically negative self-perception. This latter finding suggests that self-diminishment is an important self-perception bias in its own right (see also Noles et al., 1985) and that some of our subjects were attempting to verify their overly negative general self-views (Swann, 1983, 1987).

Thus, although our findings support the view that individuals

tend to see themselves somewhat more positively than they are seen by others, our findings also provide an important qualification to the claim that "most people regard themselves . . . as considerably more positive than is objectively likely or than reality can sustain" (Taylor, 1989, p. 6). Specifically, the overall sample trend toward self-enhancement fails to characterize either the substantial number of individuals who were quite accurate in their self-evaluations or the nontrivial number of subjects who actually derogated their performance. These findings lead us to reiterate Lewin's (1935) recommendation that psychological data be analyzed both at the aggregate level (e.g., means) and at the individual level (e.g., percentages) in order to understand both the general laws governing behavior and the individual variability that may qualify these laws and help elucidate the underlying mechanisms.

The individual variability we found in self-perception biases was related to narcissism, and this effect held for two self-report and two observer-based measures. The narcissism effect was considerably stronger than the general self-enhancement effect. Individual differences in narcissism accounted for 20% of the variance in self-perception bias, compared with the 1% explained by the self-enhancement effect. Moreover, the narcissism effect was replicated in 93% of the discussion groups, whereas the self-enhancement main effect held in only 53% of the groups.

Can the narcissism effect help explain our finding that judgments about the self were less accurate than judgments about peers? If, as we have shown, some individuals engage in self-enhancement and others in self-diminishment, then the convergence correlations between the self-judgments and the accuracy criterion would be lowered. The relation between narcissism and self-enhancement shows that there is systematic error, or method variance, in the self-judgments: A positive self-evaluation could be either a valid self-perception or narcissistic self-aggrandizement, whereas a negative self-evaluation could be either valid or unrealistic self-diminishment. In contrast, judgments about the other participants are less likely to be influenced by self-perception biases. Thus, all other things being equal, judgments about others will be more accurate than judgments about the self (cf. John & Robins, 1993).

What might be the processes and mechanisms underlying the link between narcissism and self-enhancement bias? One interpretation is based on the assumption that narcissists are more defensive than other individuals because their inflated sense of self-importance and superior competence is more easily threatened (e.g., Westen, 1990). In the present research, performing poorly in the group discussion task is likely to threaten subjects' self-image as successful future managers. Thus, when judging their performance in such an ego-involving context, narcissistic individuals should be particularly prone to experience the situation as a threat to their self-worth and attempt to bolster their self-image by positively distorting their self-perceptions.

On the other hand, when the situation and the behavior judged are not ego involving, we would not expect narcissism to influence self-perception through self-enhancement and self-diminishment biases, and self-perceptions should be no less accurate than perceptions of others. Consistent with this hypothesis, John and Robins (1993) found that self-peer convergence

was just as high as peer–peer convergence for personality traits that are affectively neutral and less ego involving.

A second interpretation is based on the finding that narcissists are interpersonally exploitative and socially inconsiderate (Millon, 1990, p. 357) and therefore likely to antagonize others in group situations. Because our group discussion task requires not only assertiveness but also working with others toward a consensual solution, the narcissists' negative interpersonal style may have compromised their effectiveness in the group.<sup>8</sup> Nonetheless, they may have evaluated their performance positively because they focused on their dominance and involvement in the task rather than on their effectiveness at accomplishing the objectives of the task; that is, they may have adopted a self-serving standard for evaluating themselves (cf. Dunning & Cohen, 1992). Future research should examine the specific behaviors narcissistic individuals exhibit in group interactions, the reactions of others to those behaviors, and the specific standards narcissists use to evaluate their own performance and that of others.

A third explanation involves the potential link between narcissism and generally positive self-views, or self-esteem. According to Freud (1914/1953) and more recent theorists (e.g., Kohut, 1971), narcissism is involved in self-esteem regulation because narcissists are constantly striving to confirm their idealized and grandiose self-views. However, the relation between narcissism and self-esteem is theoretically complex. According to clinical accounts, relatively narcissistic individuals claim high levels of self-regard, but they also have underlying feelings of inadequacy and self-doubt (e.g., Millon, 1990). Some researchers have made a conceptual distinction between genuine and defensive self-esteem, arguing that feelings of omnipotence and grandiose self-views are defenses against depression (see Raskin et al., 1991a). Consequently, most self-report measures of self-esteem are susceptible to narcissistic self-enhancement, making it difficult to interpret correlations between self-esteem and narcissism.

Empirically, the correlations between self-reported narcissism and self-esteem tend to be positive. For example, Raskin et al. (1991a) reported correlations of the NPI with several self-esteem scales, five of which were published. The correlations with these five scales ranged from .14 to .35 across three samples, with a mean of .26. Similarly, in a recent sample of 492 undergraduates, the correlation between the NPI and Rosenberg's (1965) Self-Esteem Scale was .25 (Robins, Roberts, & Covington, 1993). The item content of the NPI and the Rosenberg Self-Esteem Scale illustrates the conceptual distinctiveness of the two constructs. Whereas high self-esteem involves seeing oneself as "a person of worth, at least on an equal basis with others," narcissism involves feeling superior to others. For example, high scorers on the NPI describe themselves as special, extraordinary people who are better and more deserving than others and who feel entitled to manipulate others in an egotistical manner; in contrast, low scorers describe themselves as modest individuals who feel little need to stand out and who see themselves as no better and no worse than most people.

Overall, the relation between self-reports of narcissism and self-esteem is small and unlikely to account for the relations reported here, particularly because our findings replicated

across both self-report and observer-based measures of narcissism. Conversely, it is also unlikely that narcissism can account for the effects of self-esteem on self-enhancement processes found in previous research. Nonetheless, further research is needed to test the independent effects of narcissism and self-esteem.

### *Toward an Integration of the Correspondence and Distortion Views of Self-Perception*

We began this article by contrasting the correspondence view and the distortion view of self-perception. Rather than favoring one view or the other, our findings showed two independent components in self-perception: Although self-perceptions corresponded to a significant degree with the perceptions by others, our findings also support the conclusion that the self is a "special" kind of observer. Thus, whereas Bem's (1972) self-perception theory highlighted the similarity between self and observer perceptions, the present research suggests conditions under which veridical self-perception can be impeded by motivational processes. Our findings demonstrate that, although self-perceptions reflect observable reality, they are not fully constrained by it (Jussim, 1991).

The view that self-perceptions generally correspond with perceptions by others has served as a theoretical basis for the use of self-reports as data in psychological research. In general, our findings provide further evidence for the construct validity of self-reports (e.g., Cheek, 1982; Kenrick & Funder, 1988; McCrae & Costa, 1989). However, validity coefficients seldom reach the boundaries imposed by reliability, suggesting that nonvalid factors influence self-reports. These other components of self-report variance, typically referred to by the summary label *method variance* (D. T. Campbell & Fiske, 1959; Loewinger, 1957), are not yet well understood, and psychological analyses of such method effects are badly needed (e.g., Buckley, Cote, & Comstock, 1990; Ozer, 1989).

Although we have examined self-evaluations of performance in a specific task, rather than self-reports of global and stable traits typically assessed in personality research, our findings are relevant to the interpretation of method effects. In particular, our index of self-enhancement bias represents method-specific variance in self-perceptions of performance (i.e., variance not shared with peers or assessment staff). The correlation of this index with narcissism shows that conceptually meaningful and predictable individual differences underlie what has traditionally been conceptualized as error variance in self-reports. We have thus delineated in substantive and psychological terms a

<sup>8</sup> As a control analysis, we tested whether the link between narcissism and self-enhancement bias was due to the influence of likability. We partialled the staff's rating of likability from the correlation between narcissism and self-enhancement bias; the correlation remained significant for all four measures of narcissism. For the overall narcissism composite, the partial correlation was .44, virtually unchanged from the zero-order correlation of .46 (see Table 5). Thus, it is not the case that the staff rated the narcissists' performance more negatively just because they did not like them.

source of variance that makes self-reports different from reports by others.

The view that self-perceptions correspond with other-perceptions and the view that self-perceptions are positively distorted are not mutually exclusive. By integrating the correspondence and distortion views, we can account for both validity and bias in self-perception. Future research should further explore the boundary conditions for self-perception biases and for the lower accuracy of perceptions of self than perceptions of others. Possible factors include individual differences, characteristics of the task and the attribute judged, as well as the particular accuracy criterion and evaluation standards against which the self-judgments are assessed.

In the past, some accounts of self-perception have adopted the distortion view, emphasizing bias as a general law. Other accounts have adopted the correspondence view, emphasizing individual differences and relegating general biases to the umbrella concept of method variance. However, as Kurt Lewin (1946) emphasized

General laws and individual differences are merely two aspects of one problem: they are mutually dependent on each other and the study of the one cannot proceed without the study of the other (p. 794).

## References

- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.
- American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.). Washington, DC: Author.
- Baumeister, R. F. (1982). Self-esteem, self-presentation, and future interaction: A dilemma of reputation. *Journal of Personality*, 50, 29–45.
- Baumgardner, A. H., Kaufman, C. M., & Levy, P. (1989). Regulating affect interpersonally: When low self-esteem leads to greater enhancement. *Journal of Personality and Social Psychology*, 56, 907–921.
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). San Diego, CA: Academic Press.
- Block, J. (1978). *The Q-sort method in personality assessment and psychiatric research*. Palo Alto, CA: Consulting Psychologists Press. (Original work published 1961)
- Buckley, M. R., Cote, J. A., & Comstock, S. M. (1990). Measurement error in the behavioral sciences: The case of personality/attitude research. *Educational and Psychological Measurement*, 50, 447–474.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Campbell, J. D. (1986). Similarity and uniqueness: The effects of attribute type, relevance, and individual differences in self-esteem and depression. *Journal of Personality and Social Psychology*, 50, 281–294.
- Campbell, J. D., & Fehr, B. (1990). Self-esteem and perceptions of conveyed impressions: Is negative affectivity associated with greater realism? *Journal of Personality and Social Psychology*, 58, 122–133.
- Cheek, J. M. (1982). Aggregation, moderator variables, and the validity of personality tests: A peer-rating study. *Journal of Personality and Social Psychology*, 43, 1254–1269.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral analyses* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cooley, C. H. (1902). *Human nature and the social order*. New York: Scribner.
- Coyne, J. C., & Gotlib, I. H. (1983). The role of cognition in depression: A critical reappraisal. *Psychological Bulletin*, 94, 472–505.
- Cronbach, L. J. (1955). Processes affecting scores on “understanding of others” and “assumed similarity.” *Psychological Bulletin*, 52, 177–193.
- Cronbach, L. J., & Furby, L. (1970). How should we measure “change”—or should we? *Psychological Bulletin*, 74, 68–80.
- Dunning, D., & Cohen, G. L. (1992). Egocentric definitions of traits and abilities in social judgment. *Journal of Personality and Social Psychology*, 63, 341–355.
- Emmons, R. A. (1987). Narcissism: Theory and measurement. *Journal of Personality and Social Psychology*, 52, 11–17.
- Freud, S. (1953). On narcissism: An introduction. In J. Strachey (Ed.), *The standard edition of the complete psychological works* (Vol. 14, pp. 69–102). London: Hogarth Press. (Original work published 1914)
- Funder, D. C. (1980). On seeing ourselves as others see us: Self-other agreement and discrepancy in personality ratings. *Journal of Personality and Social Psychology*, 48, 473–493.
- Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. *Psychological Bulletin*, 101, 75–90.
- Funder, D. C. (1989). Accuracy in personality judgment and the dancing bear. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 210–223). New York: Springer-Verlag.
- Funder, D. C. (1990). Process versus content in the study of judgmental accuracy. *Psychological Inquiry*, 1, 207–209.
- 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.
- Gotlib, I. H., & Meltzer, S. J. (1987). Depression and the perception of social skill in dyadic interaction. *Cognitive Therapy and Research*, 11, 41–54.
- Gough, H. G. (1987). *The California Psychological Inventory administrator's guide*. Palo Alto, CA: Consulting Psychologists Press.
- Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*, 35, 603–618.
- Greenwald, A. G. (1982). Ego-task analysis: An integration of research on ego-involvement and self-awareness. In A. H. Hastorf & A. M. Isen (Eds.), *Cognitive social psychology* (pp. 109–147). New York: Elsevier/North-Holland.
- Greenwald, A. G., & Pratkanis, A. R. (1984). The self. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 129–178). Hillsdale, NJ: Erlbaum.
- Howard, A., & Bray, D. W. (1988). *Managerial lives in transition: Advancing age and changing times*. New York: Guilford Press.
- James, W. (1890). *Principles of psychology*. New York: Holt.
- 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.
- Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). Morristown, NJ: General Learning Press.
- Jussim, L. (1991). Social perception and social reality: A reflection-construction model. *Psychological Review*, 98, 54–73.
- Kane, J. S., & Lawler, E. E. (1978). Methods of peer assessment. *Psychological Bulletin*, 85, 555–586.
- Kenny, D. A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review*, 98, 155–163.

- Kenrick, D. T., & Funder, D. C. (1988). Profiting from controversy: Lessons from the person-situation debate. *American Psychologist*, *43*, 23-34.
- Kernberg, O. F. (1975). *Borderline conditions and pathological narcissism*. New York: Jason Aronson.
- Kohut, H. (1971). *The analysis of self*. New York: International Universities Press.
- Kruglanski, A. W. (1989). The psychology of being "right": On the problem of accuracy in social perception and cognition. *Psychological Bulletin*, *106*, 395-409.
- Kruglanski, A. W. (1993). Social constructionism and experimental social psychology: Why don't the cousins kiss? *Dialogue*, *8*, 4-7.
- Lewin, K. (1935). *A dynamic theory of personality*. New York: McGraw-Hill.
- Lewin, K. (1946). Behavior and development as a function of the total situation. In L. Carmichael (Ed.), *Manual of child psychology* (pp. 918-970). New York: Wiley.
- Lewinsohn, P. M., Mischel, W., Chaplin, W., & Barton, R. (1980). Social competence and depression: The role of illusory self-perceptions. *Journal of Abnormal Psychology*, *89*, 203-212.
- Lewis, M., & Brooks-Gunn, J. (1979). *Social cognition and the acquisition of self*. New York: Plenum Press.
- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, *3*, 635-694.
- McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology*, *43*, 293-303.
- McCrae, R. R., & Costa, P. T., Jr. (1989). Different points of view: Self-reports and ratings in the assessment of personality. In J. P. Forgas & J. M. Innes (Eds.), *Recent advances in social psychology: An international perspective* (pp. 429-439). New York: Elsevier Science.
- Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.
- Miller, D. T. (1976). Ego-involvement and attributions for success and failure. *Journal of Personality and Social Psychology*, *34*, 901-906.
- Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, *82*, 213-225.
- Millon, T. (1981). *Disorders of personality: DSM-III, Axis II*. New York: Wiley.
- Millon, T. (1990). The disorders of personality. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 339-370). New York: Guilford Press.
- Noles, S. W., Cash, T. F., & Winstead, B. A. (1985). Body image, physical attractiveness, and depression. *Journal of Consulting and Clinical Psychology*, *53*, 88-94.
- Ozer, D. (1989). Construct validity in personality assessment. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging issues* (pp. 224-234). New York: Springer-Verlag.
- Paulhus, D. L. (1990). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. Wrightsman (Eds.), *Measures of personality and social-psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.
- Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, *60*, 307-317.
- Raskin, R., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports*, *45*, 590.
- Raskin, R., Novacek, J., & Hogan, R. (1991a). Narcissism, self-esteem, and defensive self-enhancement. *Journal of Personality*, *59*, 19-38.
- Raskin, R., Novacek, J., & Hogan, R. (1991b). Narcissistic self-esteem management. *Journal of Personality and Social Psychology*, *60*, 911-918.
- Raskin, R., & Terry, H. (1988). A principal-components analysis of the Narcissistic Personality Inventory and some further evidence of its construct validity. *Journal of Personality and Social Psychology*, *54*, 890-902.
- Robins, R. W., Roberts, B. W., & Covington, M. V. (1993). Unpublished raw data, Department of Psychology, University of California, Berkeley.
- Rogers, C. R. (1959). A theory of therapy, personality, and interpersonal relations, developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 3, pp. 185-256). New York: McGraw-Hill.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Swann, W. B., Jr. (1983). Self-verification: Bringing social reality into harmony with the self. In J. Suls & A. G. Greenwald (Eds.), *Psychological perspectives on the self*. Hillsdale, NJ: Erlbaum.
- Swann, W. B., Jr. (1987). Identity negotiation: Where two roads meet. *Journal of Personality and Social Psychology*, *53*, 1038-1051.
- Taylor, S. E. (1989). *Positive illusions: Creative self-deceptions and the healthy mind*. New York: Basic Books.
- Taylor, S. E., & Brown, J. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193-210.
- Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181-227). San Diego, CA: Academic Press.
- Tesser, A., & Campbell, J. D. (1982). Self-evaluation maintenance and the perception of friends and strangers. *Journal of Personality*, *50*, 261-279.
- Thorton, G. C., & Byham, W. C. (1982). *Assessment centers and managerial performance*. San Diego, CA: Academic Press.
- Westen, D. (1990). The relations among narcissism, egocentrism, self-concept, and self-esteem: Experimental, clinical, and theoretical considerations. *Psychoanalysis and Contemporary Thought*, *13*, 183-239.
- Wink, P. (1991). Self- and object-directedness in adult women. *Journal of Personality*, *59*, 769-791.
- Wink, P., & Gough, H. G. (1990). New narcissism scales for the California Psychological Inventory and MMPI. *Journal of Personality Assessment*, *54*, 446-462.

Received August 13, 1991

Revision received June 18, 1993

Accepted June 22, 1993 ■