Journal of Experimental Psychology: General
2015, Vol. 144, No. 5, 1003–1020

0096-3445/15/$12.00 http://dx.doi.org/10.1037/xge0000105

Self-Enhancement Diminished

Patrick R. Heck and Joachim I. Krueger
Brown University

Self-enhancement is a positive bias in self-perception, which may imply error. However, conventional measures of self-enhancement are difference scores that do not distinguish a positive bias from a self-enhancement error, that is, they fail to identify those individuals who hold an irrationally or inaccurately positive view of themselves. We propose 2 new measures to separate error from bias. In the domain of personality judgment, we estimate a defensible bias and an enhancement error from individuals' actual and perceived similarity with others. In the domain of performance, we adapt a decision-theoretic framework to distinguish those who falsely believe to be better than average from those who actually are better. We illustrate the properties of these measures in 3 empirical studies and computer simulations. Implausibly high majorities of people consider themselves to be above average on various dimensions.

Keywords: self-enhancement, accuracy, judgment and decision-making, measurement

> "Implausibly high majorities of people consider themselves to be above average on various dimensions" (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001, p. 348).

Self-enhancement is a widely studied phenomenon in social perception. Its empirical existence is rarely contested, particularly in the context of Western, individualist cultures (Alicke & Sedikides, 2011). Svenson (1981) famously reported that most people think they are better drivers than average, and nearly all college professors think they are better teachers than their colleagues (cf. Dunning, 2012). Because somebody must be worse than average if there is any variation at all, some positive self-assessments must be wrong. The personal and social costs of self-enhancement seem self-evident. Individuals who perceive themselves irrationally and inaccurately make poor decisions at some point (Larrick, Burson, & Soll, 2007; Moore, Kurtzberg, Fox, & Bazerman, 1999) or are rejected by their peers (Hoorens, Pandelaere, Oldersma, & Sedikides, 2012; Paulhus, 1998; Robins & Beer, 2001). However, self-enhancement has its appeal because it has direct hedonic benefits and because it may contribute to other aspects of well-being (Taylor & Brown, 1988).

Assured of the robustness of self-enhancement, researchers began to explore its underlying psychological processes and variations in strength (Sedikides & Gregg, 2008). Cultural context, the type of judgment task, and method of measurement have emerged as important moderators. With regard to culture, an influential claim is that individualists self-enhance more than collectivists do (Heine & Hamamura, 2007; see Brown, 2010, for a critique). With regard to task, self-enhancement is stronger for matters of morality or interpersonal warmth than it is for issues of intelligence or competence (Allison, Messick, & Goethals, 1989; Brambilla & Leach, 2014; Paulhus & John, 1998). Within the performance domain, self-enhancement emerges for easy tasks or tasks of moderate difficulty, whereas difficult tasks can elicit self-effacement (Kruger, 1999). Some of these moderator effects play a role in the debate over whether processes of information processing are sufficient to explain self-enhancement (Chambers & Windschitl, 2004; Moore & Healy, 2008), or whether self-enhancing motives must be assumed to operate (Alicke, 1985; Brown, 1986, 2012).

We begin with a review of prominent theoretical perspectives and the measurement indices they have produced (see also Krueger & Wright, 2011; Kurt & Paulhus, 2007). We show that these indices conflate a false sense of superiority with true superiority. The view that the proportion of those who consider themselves to be above average is "implausibly high" (Baumeister et al., 2001, p. 348) raises the question of how low the proportion and how accurate an individual's judgments should be to be considered adequate. Building on research on accuracy and bias in behavioral predictions (Epley & Dunning, 2000, 2006), knowledge claiming (Paulhus, Harms, Bruce, & Lysy, 2003), and personality judgment (Zuckerman & Knee, 1996), we develop two new measures that separate self-enhancement *error* from rationally defensible *bias*. The new measures are therefore more conservative with respect to attributions of irrationality.

Our theoretical perspective draws on the logic of statistical regression and standard decision theory. The first method of measurement is applicable to a variety of judgment domains that have no clear reality criterion. We study personality judgment to demonstrate the properties of this method. The second method is applicable when objective, external criteria are available. We study test performance to demonstrate the properties of this method. In three studies—each accompanied by a computer simulation—we

Patrick R. Heck and Joachim I. Krueger, Department of Cognitive, Linguistic & Psychological Sciences, Brown University.

We are grateful to Joseph Austerweil for thoughtful comments on a draft of this article and to Don Moore for generously providing test materials for Studies 2 and 3.

Correspondence concerning this article should be addressed to Patrick R. Heck or Joachim I. Krueger, Department of Cognitive, Linguistic & Psychological Sciences, Brown University, 190 Thayer Street, Providence, RI 02912. E-mail: patrick_heck@brown.edu or joachim_krueger@brown.edu

illustrate the properties of these new indices and compare them to conventional ones.

Traditional Measures of Self-Enhancement

The most popular theoretical perspective on self-enhancement is grounded in the theory of social comparison (Festinger, 1954). According to this perspective, self-enhancement arises from active processes of self-other comparison. The other person being judged is typically taken to be the average person of a reference group. Evidence for self-enhancement then takes the form of the better-than-average effect (BTAE; Alicke, 1985; Brown, 1986). To measure the BTAE directly, respondents are asked to rate how good (smart, happy, handsome, etc.) they are relative to the average person. This direct measure poses difficulties (Krueger, Freestone, & MacInnis, 2013; Krueger & Wright, 2011). The assumption that potentially irrational individuals make separate judgments of self and other, and then reliably compress and scale the relationship between the two into a single dimension is questionable. Furthermore, this direct measure neglects important information. Once the compression is made, the constituent perceptions of self and other are lost. Although the direct measure enjoys continued use (e.g., Eriksson & Funcke, 2014), we consider it no further because of these difficulties. We turn instead to the indirect measurement of self-enhancement, which requires respondents to make separate judgments of the self, S, and the average other, O, which the investigator then converts to a difference score, S—O (Klar & Giladi, 1999).

When the self is rated higher on positive attributes and lower on negative attributes than another person, there is *prima facie* evidence of self-enhancement. The opposite differences indicate self-effacement. Only respondents who produce a difference score of precisely zero escape the charge of bias (Krueger & Funder, 2004). More important, for the present purpose, the difference-score measure does not distinguish bias from error; it does not reveal who correctly claims to be better than average and who does so mistakenly. The former type of person will be wrongly labeled a self-enhancer, while only the latter is in error.

Noting that the social comparison approach does not provide an estimate of what the person "is really like," researchers developed a social reality approach (Colvin, Block, & Funder, 1995; John & Robins, 1994). They introduced estimates of true scores, T, such as test scores or aggregated observer judgments as criterion measures. When $S > T$, the respondent overestimates his or her performance (Moore & Healy, 2008). Like the social comparison approach, the social reality approach allows only one point to represent accuracy. Epley and Dunning (2000), for example, asked participants to predict how much money they would donate to charity and then observed how much these participants actually donated. A discrepancy of 1c was sufficient for the demonstration of bias. In a later series of studies, Epley and Dunning (2006) measured bias and accuracy separately. Bias was again represented by positive difference scores between self-prediction of prosocial behavior and actual behavior, whereas accuracy was represented by profile correlations between judgments and criteria over instances. This work yielded the important result that correlational accuracy can coexist with overall bias.

Being able to estimate correlations between self-judgments, S, and criteria, T, allows investigators to move beyond group data and assess individuals. A structural property of this approach is that it produces lawful regression effects (Fiedler & Krueger, 2012; Fiedler & Unkelbach, 2014). Someone with a low true score is more likely to self-enhance than someone who has a high score (Krueger & Mueller, 2002; Moore & Small, 2007). Such regression effects make it difficult to separate unique properties of self-enhancement from the mere location of the respondent's true criterion score T (see Krueger, 1998b, for a method that avoids regression effects).

In the domain of personality judgment, the social reality approach requires a set of judgments about the target person (Funder, 1995). These are typically provided by knowledgeable others, such as mates, peers, or "expert observers." Observer judgments of the target are then aggregated and subtracted from self-judgments.[1] Aggregation is useful as it increases reliability, but it does not guarantee increases in validity. A wisdom-of-the-crowd benefit may occur, but it is far from certain (Krueger & Chen, 2014; Mannes, Soll, & Larrick, 2014). For our purposes, the critical limitation of the social reality approach is that it, like the social comparison approach, fails to separate error from bias. Like the social comparison approach, it identifies most respondents as either self-enhancers or self-effacers.

Individuals who rate themselves differently (more or less favorably) than observers rate them may indeed be accurate in their judgments. They may have access to private information others cannot see, or they may have a more complete view of themselves across social contexts (Fiedler, 1996, 2000; Krueger, Ham, & Linford, 1996; Malle & Pearce, 2001). For example, a group of observers who are all related to the target are a statistically biased sample, and they may well be psychologically biased. Sampling bias is difficult to eliminate because there are no known ways of determining a person's true life space. Lacking a defensible definition of the population of potential observers, it is not possible to sample them representatively.[2] Aggregating observer judgments dilutes random error, but with the population unknown, it is not clear whether the aggregation of the sampled judgments only allows the true signal to emerge or whether it also brings shared observer bias into focus (Kruglanski, 1989). In performance domains, the sampling challenge is more tractable because test constructors may define a population of relevant tasks (Anastasi, 1988; Haynes, Richard, & Kubany, 1995). Here, the index S—T is trustworthy inasmuch as the two variables are measured reliably.

The social comparison and the social reality approaches yielded opposite conclusions regarding the Taylor-and-Brown hypothesis. Taylor and Brown (1988) believed that self-enhancement is a cognitive illusion, but noted its association with well-being and social adjustment. Most of the findings they reviewed were obtained with the social comparison approach. In contrast, Colvin et al. (1995) and others argued that self-enhancement is maladaptive, and they offered evidence from studies adopting the social reality

[1] A variant of this method is to regress S on the aggregated observer judgments and retain the residual as an index of self-enhancement (Krueger & Wright, 2011).

[2] When a population is well defined and ready to be sampled randomly, the average observation is most informative. This is not so when a population is ill defined or not defined at all. The best observer of a target individual's personality (e.g., a spouse, friend, or therapist) is most likely not the average observer in the pool.

approach (see also Kurt & Paulhus, 2008). It is possible to reconstruct these conflicting patterns from methodological assumptions and practices, which means that neither the social comparison nor the social reality approach has been able to settle the Taylor-and-Brown debate (Krueger & Wright, 2011). We will return to the Taylor-and-Brown hypothesis and ask whether it can be tested with new measures.

Some authors have recognized the limitations of the conventional measures. Moore and Small (2007) recommended a joint display of both. When self-judgments, S, and other judgments, O, are plotted against true values, T, a robust pattern (that we replicate) emerges: low scorers overestimate their own performance (S > T) but underplace themselves relative to others (S < O). In contrast, high scorers underestimate their own performance (S < T), but overplace themselves relative to others (S > O). Moore and Small's theoretical account (which we adopt) assumes "differential regression," that is, the prediction that judgments of others are more regressive with respect to actual performance scores than are self-judgments. A casual look at the Moore and Small pattern may suggest that social comparison scores, S—O, are negatively correlated with social reality scores, S—T, but this is not so. The two difference-score measures will be positively correlated in most data sets because they have one variable in common (Krueger & Wright, 2011).

Kwan, John, Kenny, Bond, and Robins (2004) regarded the social comparison index and the social reality index to be main effects of social judgment. The social comparison index is a perceiver effect because the respondent generates the estimates of O. The social reality index is a target effect because the measure T is "about" the respondent. Kwan et al. then recast self-enhancement as the interaction of perceiver and target effects. Adapting Kenny's (1994) social relations model, they estimated the interaction term by subtracting both the respondent's average judgment of others and the average judgment made by others about him or her from the respondent's self-judgment. They then added the grand average of all judgments back in, but as this term is a constant it does not affect correlations. This hybrid index (henceforth: S—O—T), though intended to capture an interaction effect, is almost identical to the sum of the two conventional indices ([S—O] + [S—T] = 2S—O—T; Krueger & Wright, 2011). Therefore, we hypothesize that the hybrid measure shares the limitations of the earlier measures of social comparison and social reality, and we test this hypothesis in two studies and computer simulations.

Two New Measures of Self-Enhancement

We propose a novel integration of the social comparison and social reality perspectives. We ask how a rational person might be expected to make judgments about others (Proposal 1) and how social judgments and external criteria can be combined to identify mistaken self-enhancers, that is, those who falsely believe they are better than average (Proposal 2). The first and the second measure are, respectively, grounded in theories of regression and decision analysis. Each measure is designed to separate those individuals who falsely claim superiority from those who do so justifiably. As a result, the new measures are more conservative than the traditional ones. We do not presuppose the existence of self-enhancement error; instead, we seek to stimulate empirical tests of self-enhancement bias and error along with assessments of rationality and accuracy. We recommend that students of self and social perception use (and improve) these measures when testing debated moderators of self-enhancement and novel theoretical approaches to the phenomenon.

The Projection-Based Index

How should one make judgments of the average person? If everyone were omniscient and rational, everyone would make the same correct judgment of the average person. If so, the S—O measure would be perfectly correlated with self-description, S, thereby confounding self-enhancement with self-description. Conversely, if all individuals described themselves as average, all difference scores would be zero, and self-description would reduce to social description, or self-stereotyping. Neither self-enhancement bias, nor accurate or erroneous self-description could be detected. If the index S—O is too inclusive (i.e., by classifying many above-average individuals as self-enhancers), a measure is needed that raises the threshold for the detection of self-enhancement error and that does so by describing the process by which respondents make judgments of the average person.

To construct such a measure, we tap into a century of research on social projection (Allport, 1924; Krueger, 2007). Social projection refers to the result that most people perceive others as being similar to them and to the processes that produce this result (Krueger, 2007). Projection implies that social judgments are partly egocentric, and most strongly so when others are ill-defined or depersonalized. Empirically, there is a robust positive correlation between self-judgments, S, and judgments of others, O. Hoch (1987) showed that this must be so if O judgments are to be accurate (see also Dawes, Faust, & Meehl, 1989). By necessity, most people are in the majority most of the time, a condition that justifies the projection heuristic. The more people project from themselves to others, the more accurate their judgments will be.

The use of social projection as an inference heuristic is limited when people have good information about individual others (Ames, 2004; Krueger, 1998a). In such a case, the correlation between self-judgments and social judgments may be spurious, in the sense that it is the result of actual rather than assumed similarity. Furthermore, it is possible that people infer their own attributes from sampling information obtained in their social world (e.g., Galesic, Olsson, & Rieskamp, 2012). In the research we describe below, we address domains in which the uncertainty about the attributes of others is likely greater than the uncertainty about one's own, and hence the assumption of social projection is reasonable.

Social projection facilitates judgmental accuracy overall, but it may be too strong or too weak. Hoch (1987) recommended that assumed similarity (projection) should track actual similarity, that is, the correlation between the person's self-judgments and the average self-judgments in the group. Taking Hoch's rule as a rational benchmark, we explore its implications for measures of self-enhancement. To do so, we compute predicted other judgments, P, for each respondent over a set of traits (Krueger et al., 2013). Specifically, we regress average self-judgments made by members of the group on the individual respondent's own self-judgments. This regression equation yields predictions of judgments of the average person. These predicted other judgments—

and the rational bias they imply—vary with the respondent's typicality of the group. A typical group member should predict the average group member to be similar to him- or herself, whereas an atypical member should assume greater differences. A typical group member will, therefore, have a smaller rational bias and more room for self-enhancement error. We empirically explore and mathematically simulate these regularities.

Over traits, we expect that most people's self-judgments, S, are more strongly and positively correlated with the traits' desirability, D, than are their judgments of others, O (Moore & Small, 2007). Most people claim to be kind and sincere, while few admit to being selfish (Krueger, 1998b). Because the predicted other judgments, P, are regressive (i.e., not perfectly correlated) with respect to S, they will be less strongly associated with trait desirability. This difference in slope is the basis for a rational self-enhancement bias. Individuals with positive self-images can be expected to show this bias if they project their own traits imperfectly to the group. In contrast, the question of self-enhancement *error* is addressed by comparing the association between judgments of others, O, with their respective predicted values, P. We submit that there is self-enhancement error if O judgments are less strongly related to trait desirability than are the P estimates.

As noted above, we assume individuals estimate the personality attributes of others projectively from their self-judgments. This assumption does not mean, however, that self-judgments are necessarily more accurate than other judgments. In the social comparison paradigm, self-enhancement bias may arise from inflated self-judgments or from deflated other judgments. To account for this dual source of bias, we will adapt the projection-based index and show that the separation of bias and error does not depend on assumptions about the differential accuracy in S and O.

To review, the projection-based measure acknowledges that most people have positive self-images and that they face a difficult task when judging the personality of the average person. Using their own self-judgments as cue-valid information and weighting them by the representativeness of their self-judgments for the group, they can generate reasonable predictions. The logic of regression entails a self-enhancement bias. Therefore, the measure of self-enhancement *error* must identify respondents who rate the average person less positively than their own rational predictions permit. In Study 1, we illustrate the use of the projection-based measure. We identify the respondents who show self-enhancement according to the conventional social-comparison measure, and then separately estimate the proportions of rational bias and self-enhancement error. We expect to find that the conventional metric overestimates the prevalence of self-enhancement as an error.

The Decision-Theoretic Index

The projection-based index of self-enhancement is suitable for analyses over traits in perceptions of personality. The index does not comprise a reality criterion, but is grounded in assumptions about rational judgment in an uncertain environment or environments that do not provide reality criteria. To measure self-enhancement in domains of competence and performance, we turn to test scores as estimates of true scores, T, and use them to integrate the social comparison perspective with social reality.

We define self-enhancement error as the false belief of being more competent than the average person. With this definition, those individuals who correctly claim to be better than average will not be charged with self-enhancement *error*. To identify error, three sets of data are needed. Performance data must be scored to yield an index T representing the criterion. Then, respondents provide S and O estimates as they do in the social comparison paradigm. With these data, respondents can be categorized into four groups. First, respondents who believe they performed better than average (median), $S > O$, and did, $T > \bar{T}$, are Hits, H. Second, respondents who believe they scored better than average but did not, are False Alarms, FA. Third, respondents who believe they scored like the average person or worse but did not, are Misses, M. Fourth, respondents who believe they scored like the average person or worse and did, are Correct Rejections, CR.

Use of these four types casts the respondents' task to make S and O judgments as a statistical decision problem (Swets, Dawes, & Monahan, 2000). Individuals motivated by accuracy will seek to maximize the chances of being an H or a CR. This is difficult because T and $\bar{T}$ are unknown. Making judgments under uncertainty, respondents must consider how much they care about making either type of error; they need to know what it means to them to be prudent (Pascal, 1669/1962). Respondents who are more troubled by the possibility of missing a positive result (a Type II error) than by falsely claiming one (Type I error) will tolerate a positive decision bias. They will be more liberal in letting their S judgments be greater than their O judgments than the reverse (Haselton, Bryant et al., 2009; Lynn & Feldman Barrett, 2014).

Paulhus, Harms, Bruce, and Lysy (2003) developed a decision-theoretic index of self-enhancement for knowledge domains. Individuals who claim knowledge they cannot have are self-enhancers or "overclaimers." The overclaiming questionnaire (OCQ) comprises 150 items (e.g., "Manhattan Project"), each of which respondents rate on a scale from 0 (*never heard of it*) to 6 (*very familiar*). Twenty percent of the items are foils ("cholarine"), however, so that any claim of knowledge is false. The overclaiming measure is a decision-theoretic variant of the social reality approach. It is not sensitive to social comparison and respondents' sense that they are better or worse than the average person. The index we propose combines essential features of the social reality approach with those of the social comparison approach.[3]

Epley and Dunning (2000, 2006) also took a decision-theoretic perspective when studying respondents' predictions of their own behavior. These authors distinguished between mean-level prediction errors (bias) and discrimination accuracy over items. We build on this work by separating bias from accuracy for individual judgments, while also examining discrimination accuracy using correlational measures. Finally, we demonstrate differences in discrimination accuracy between accurate and errant self-enhancers.

We designed Studies 2 and 3 to study self-enhancement bias and error in a performance domain. We collected S, O, and T data to classify each respondent into one of the four decision-theoretic categories and asked whether the conventional scores (S—O,

[3] Overclaimed knowledge is a decision-theoretic False Alarm. To validate their measure of self-enhancement, Paulhus et al. used the aggregate of Hits and False Alarms, however, and thereby focused on bias, not on error.

S—T, and S—O—T) overestimate the prevalence of self-enhancement error. The overdiagnosis of self-enhancement is of theoretical and practical interest because it implies a reverse-inference fallacy (Krueger, 2014; Wason, 1960). In propositional logic, a reverse inference is invalid. If it is true that all men are mortal, it does not follow that a mortal being is a man. In a probabilistic environment, a situation may arise in which the probability of a particular condition, such as a physical disease or a psychological problem, is lower than the probability of the diagnostic sign that is associated with that condition (Bar-Hillel, 1980). The probability of the diagnostic sign given that the condition is present is higher than the probability of the condition given the sign. The context of self-enhancement presents an analogous asymmetry: conventional measures treat both Hits and False Alarms as cases of self-enhancement; that is, they assume an identity of sign and condition. If, however, it is recognized that a Hit is not a judgmental error, it can be seen that the probability of self-enhancement error (i.e., a False Alarm) given a positive conventional index is higher than the probability of the error given a positive index. Error implies bias, but bias does not entail error. In Study 3, we provided some respondents with correct feedback about their performance, thereby constraining their options for strategic self-enhancement.

The projection-based and decision-theoretic measures of self-enhancement may yield new insights into the interplay of rational and mistaken judgment. They do so by preserving the theoretical legitimacy and importance of both the social comparison and the social reality perspectives. In three empirical studies—each supplemented with a computer simulation—we uncouple genuine judgment error from rational and accurate judgment processes and biases. The contribution of this work is theoretical and methodological; indeed, the two are closely intertwined (Greenwald, 2012). Our approach does not prejudge the outcome of further research on the motivational and cognitive bases of self-enhancement, its cultural variability, or its social or biological adaptiveness. Instead, the proposed indices build on previous work (e.g., Epley & Dunning, 2006; Paulhus et al., 2003; Zuckerman & Knee, 1996) to create opportunities to revisit these questions with renewed vigor and purpose.

Study 1: Personality

Method

Participants and materials. Students ($N = 83$; 63 female, mean age = 19.48 years) in a personality psychology course at Brown University participated as volunteers. Two individuals provided incomplete data, leaving an effective sample size of 81. The data of another seven participants were excluded from the correlational and regression analyses for lack of variation in ratings of "the average person."

A questionnaire was compiled with 10 positive and 10 negative trait adjectives from Otten and Wentura (2001). One negative trait was not in the final survey because of experimenter error. Table 1 shows the remaining trait adjectives in ascending order of social desirability. Participants rated each trait on a scale ranging from 1 (*not at all*) to 5 (*extremely*) in response to the following prompts presented in counterbalanced order: "Please rate how well the following traits describe you," "rate how well the following traits describe the average person," and "rate how socially desirable the following traits are." The order of the trait adjectives was randomized once and then maintained for all presentations. On the last page of the survey, participants were asked to indicate their gender, age, and academic standing.

Table 1
Mean Trait Adjective Ratings of Self, Others, and Social Desirability

Trait adjective	Self	Other	Desirability
Unfair	1.80 (.63)	2.46 (.65)	1.02 (.16)
Unfriendly	1.43 (.57)	2.30 (.56)	1.17 (.38)
Cold	1.64 (.64)	2.31 (.58)	1.22 (.45)
Cowardly	2.10 (.75)	2.51 (.71)	1.23 (.46)
Boring	1.90 (.76)	2.33 (.65)	1.26 (.49)
Sad	2.13 (.80)	2.30 (.53)	1.43 (.57)
Pessimistic	2.23 (.81)	2.53 (.65)	1.44 (.57)
Hectic	2.58 (1.04)	2.52 (.78)	1.53 (.65)
Passive	2.55 (.94)	2.71 (.61)	1.83 (.77)
Energetic	3.39 (.90)	2.86 (.61)	3.63 (.81)
Creative	3.46 (1.00)	2.57 (.63)	3.85 (.95)
Strong	3.41 (.91)	2.84 (.69)	3.90 (.82)
Self-confident	3.19 (.86)	2.90 (.66)	4.06 (.78)
Empathetic	3.99 (.88)	2.92 (.78)	4.17 (.76)
Happy	3.60 (.83)	3.17 (.66)	4.26 (.74)
Sociable	3.55 (.79)	3.14 (.59)	4.33 (.63)
Open-minded	3.93 (.73)	2.80 (.69)	4.33 (.61)
Friendly	3.92 (.69)	3.20 (.56)	4.59 (.63)
Reliable	4.19 (.72)	2.75 (.67)	4.59 (.59)
Mean	2.89	2.69	2.84
SD	.90	.30	1.47
SE	.21	.07	.34

Note. All ratings were made on a Likert scale ranging from 1 (*not at all*) to 5 (*extremely*). Traits are sorted in ascending order of desirability. $N = 81$.

Procedure. Participants were asked to sit quietly for the duration of the study. Upon completion of the ratings, participants returned the questionnaires to the experimenter who then provided a debriefing.

Results

Descriptives. Table 1 shows the means of the self-judgments, S, the judgments of the average person, O, and the judgments of trait desirability, D. These three sets of averages were positively intercorrelated. Self-judgments, $r_{SD} = .96$, and other judgments, $r_{OD} = .86$, were both closely associated with trait desirability, and self-judgments were positively correlated with other judgments, $r_{SO} = .83$. The size of these correlations reflects the effect of aggregation, which filters out idiosyncratic biases and errors. In the following, we focus on the associations among the three rating variables at the individual level.

Projection-based self-enhancement. How much self-enhancement may emerge from rational thinking? We propose the following logic: If a person knows—or guesses well—the correlation between self-judgments and the traits' mean self-judgments, then this person can generate rational estimates of the average person by taking his or her own self-judgments and weighting them by that correlation (Krueger et al., 2013). To model this strategy with our data, we regressed average self-judgments onto individual self-judgments over all respondents and traits. For each respondent and trait, we then computed a predicted judgment of the average person as $p = 1.32 + .546 \times$ self-judgment.

Table 2 shows the intercorrelations of S, O, D, and P computed over respondents. We first note the familiar BTAE. Trait desirability judgments predicted self-judgments better (*mean b* = .59) than they predicted other judgments (*mean b* = .18), $t(73) = 11.80$, $d = 1.73$. This regression-based measure of the BTAE (i.e., the difference between the two regression weights) is equivalent to the traditional difference-score index. To demonstrate this equivalence, we computed a difference-score index by subtracting the self-judgments on the negative traits from the average self-judgments on the positive traits for each respondent ($M = 1.60$, $SE = .08$) and subtracting the corresponding difference score obtained with other judgments ($M = .54$, $SE = .07$). The mean difference was significant, $t(73) = 10.79$, $p < .001$, and the effect size was large, $d = 1.59$. Our regression index of the BTAE ($b_{SD} - b_{\mathrm{OD}}$) was almost perfectly correlated with this traditional difference-score index, $r(72) = .995$, a finding that is theoretically reassuring, but psychometrically unsurprising.

Next, we found the expected evidence for social projection, in that self-judgments predicted other judgments, although the effect size was small (*mean b* = .17). This measure of social projection is correlational, which leaves open the possibility that the association between self- and other judgments reflect, in part, processes of self-stereotyping (Cho & Knowles, 2013; Van Veelen, Otten, Cadinu, & Hansen, in press). Having computed projection-based predictions P as normative benchmarks for judgments of the average other, we found that trait desirability ratings predicted self-judgments more strongly (*mean b* = .58) than they predicted projection-based other judgments (*mean b* = .32), $t(73) = 15.48$, $p < .001$, $d = 1.23$. This difference ($b_{SD} - b_{\mathrm{PD}}$) reflects the rational component of self-enhancement bias. To test for *self-enhancement error*, we asked whether the empirically observed other judgments were less strongly associated with trait desirability than were the projection-based predictions ($b_{\mathrm{PD}} - b_{\mathrm{OD}}$). Indeed they were, $t(73) = 5.89$, $p < .0001$, $d = .86$. Figure 1 displays the whole pattern by showing the decreasing slopes of self-judgments, predicted other judgments, and obtained other judgments against trait desirability.

Almost all respondents (93%) self-enhanced by the conventional standard, $b_{SD} > b_{\mathrm{OD}}$, and a similarly high percentage (97%) showed a rational, projection-based bias, $b_{SD} > b_{\mathrm{PD}}$. Only 77% showed self-enhancement *error*, $b_{\mathrm{PD}} > b_{\mathrm{OD}}$. This is still a large section of the sample, but, it is clearly smaller than the conventional section. There is a sizable minority of individuals (23%) who self-enhance without error. This pattern points to the danger of reverse inference. While every respondent who was identified as showing self-enhancement error by our measure also showed self-enhancement bias by the conventional measure, the inverse was not true. Of those who conventionally self-enhanced, only 83% went beyond the rationally defensible benchmark. Tables 3 through Table 7 shows the crossed frequencies for conventional bias and true error from which these conditional probabilities were computed.

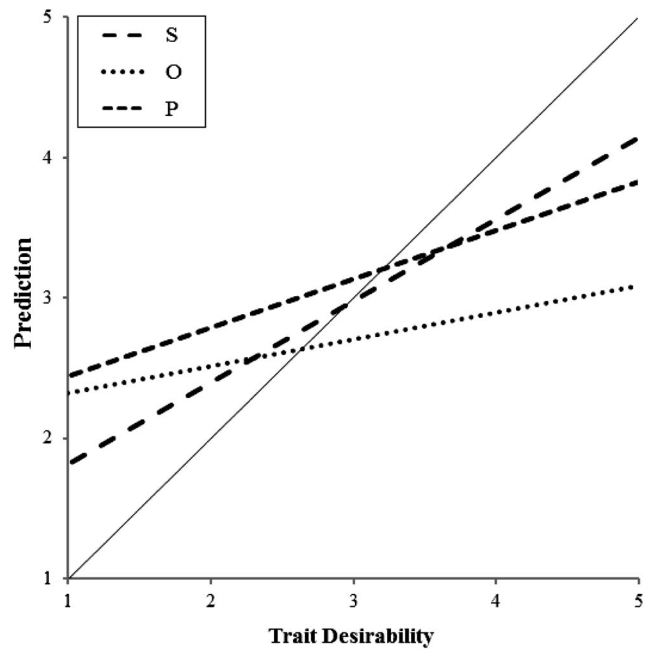

Figure 1. Predicted values generated by regressing self-judgments S, judgments of the average person O, and predicted judgments of the average person P on trait desirability (Study 1).

Individual differences. When predicting a rational person's judgments of the average other, we used the correlation over all self-judgments and group averages. We assumed that there is one general association that an idealized social perceiver might know and use to generate estimates of the average person. A more fine-grained approach, as envisioned by Hoch (1987), is idiographic. Therefore, we extended our analyses by predicting each respondent's other judgments from his or her self-judgments and the correlation between this respondent's self-judgments and the average self-judgments obtained from the group of respondents.

Table 2
Descriptive Statistics for Primary Correlations and Regressions (Study 1)

	Z score			Slope		Intercept	
	M	M_r	*SD*	*M*	*SD*	*M*	*SD*
r_{SD}	1.00	.76	.42	.58	.24	1.24	.68
r_{OD}	.59	.46	.56	.19	.21	2.13	.58
r_{SO}	.35	.34	.50	.17	.24	2.17	.70
r_{SP}	—	1.00	0	.55	0	1.32	0
r_{PD}	1.00	.76	.42	.32	.13	2.00	.38
r_{PO}	.35	.34	.50	.25	.38	2.22	1.05

Note. S = self; O = other; P = prediction of other; D = desirability. Slopes and intercepts are unstandardized (*b*). $N = 74$.

Table 3
Crossed Frequencies for Conventional Bias and Self-Enhancement Error

	$b_{\mathrm{PD}} > b_{\mathrm{OD}}$	$b_{\mathrm{PD}} \leq b_{\mathrm{OD}}$	Total
$b_{SD} > b_{\mathrm{OD}}$	57	12	69
$b_{SD} \leq b_{\mathrm{OD}}$	0	5	5
Total	57	17	$N = 74$

Table 4
Crossed Frequencies for Social Comparison Index and True Error

	False Alarm	~(False Alarm)	Total
S—O > 0	44	79	123
S—O ≤ 0	0	66	66
Total	44	145	*N* = 189

Table 6
Crossed Frequencies for Hybrid Index and True Error

	False Alarm	~(False Alarm)	Total
S—O—T + $\bar{T}$ > 0	44	75	119
S—O—T + $\bar{T}$ ≤ 0	0	70	70
Total	44	145	*N* = 189

We found that, on average, this decomposition of total self-enhancement into bias and error terms was virtually identical to the simpler analysis reported here.

However, the idiographic approach offers additional opportunities for the exploration of theoretically important questions. Specifically, we can demonstrate two dependencies that are built into the projection model. The first dependency is that the more respondents project their own traits onto others, the less they will self-enhance. At the limit, when other judgments are identical to self-judgments, any social comparison index of self-enhancement will be nil. Second, the more representative a person's self-judgments are of the group, the more similar projection-based other judgments are to self-judgments. Rational, projection-based predictions of others' personalities will then yield little self-enhancement, which means that there will be more room for self-enhancement *error*. To recover these two implications from our data, we regressed the index of self-enhancement error (b_{PD} - b_{OD}) simultaneously on the respondents' projection scores and their typicality scores (when their individual S are regressed on group average S). The results corroborated our theoretical analysis. Self-enhancement error decreased with social projection, $\beta = -.74$, $t(71) = -8.99$, $p < .01$, but increased with the person's typicality of the group, $\beta = .44$, $t(71) = 5.32$, $p < .01$.

Location of bias. In the foregoing analysis, we anchored the separation of bias and error on self-judgments so that other judgments were classified as being either more or less negative than rational inference would suggest. Self-enhancement error was thereby cast as an excessive diminishment of others. As noted earlier, self-enhancement error may also arise from self-judgments being too positive. Indeed, this interpretation of self-enhancement is the prevailing one in the literature (e.g., Epley & Dunning, 2006).

To account for this perspective, we adapted the projection-based index by anchoring analysis on O judgments and reverse-predicting S judgments. In other words, we asked how positive S judgments would have been if respondents projected rationally and if they ended up with the O judgments they actually made. We rewrote the regression equation as $p = 1.32 + (O / .546)$, so that P now refers to the inferred self-judgment. We then asked whether P was less strongly associated with trait desirability than were observed S judgments. Regressing P on D, we find that mean *b*PD = .38. The inferred self-judgments were less positive than observed self-judgments (*mean bSD* = .58), $t(73) = 4.76$, $p < .01$, $d = .72$, suggesting self-enhancement error. In other words, participants viewed themselves more positively than was predicted by the normative benchmark. The difference between inferred self-judgments and observed other judgments (*mean b*OD = .19), $t(73) = 7.75$, $p < .01$ $d = .51$, indicated rational bias. Classifying individuals into rational and erroneous self-enhancers yielded the same results as our initial analysis. To repeat, the projection-based measure does not compel the view that self-judgments are more accurate than other judgments; it only assumes that the latter are, for the most part, inferred from the former. Both S and O may be biased or in error, and our method yields the same decomposition into bias and error irrespective of which variable is given and which is inferred.

Table 5
Crossed Frequencies for Social Reality Index and True Error

	False Alarm	~(False Alarm)	Total
S—T > 0	34	55	89
S—T ≤ 0	10	90	100
Total	44	145	*N* = 189

Simulating Enhancement Bias and Error

To map out the differential consequences of projection and typicality for self-enhancement error, we conducted computer simulations in MATLAB (Mathworks, 2013). In one simulation, we set the means for social projection (r_{SO} = .3) and self-group typicality (r_{SmeanS} = .7) as well as their *SD*s (*SD* = .2) to values reflecting the empirical findings. We varied self-positivity, r_{SD}, from −1 to 1, and drew 10,000 observations at each level of self-positivity. We computed estimates of other positivity, r_{OD}, as the product of self-positivity and social projection, and computed predicted other positivity, r_{PD}, as the product of self-positivity and typicality (see DiDonato, Ullrich, & Krueger, 2011, for the logic of this estimation rule). The results (see Figure 2, Panel a) show that there is self-enhancement error ($r_{OD} < r_{PD}$) for all positive self-images ($r_{SD} > 0$). In a second simulation, we reversed the mean correlations so that projection (.7) exceeded typicality (.3). The

Table 7
Descriptive Statistics and Intercorrelations (Study 2)

Measures	*M* (*SD*)	Correlation O	T	S—O	S—T	K
Self-judgment (S)	12.39 (4.01)	.20	.69	.79	.67	.53
Other-judgment (O)	10.78 (2.79)	—	.04	.32	.23	−.15
Total score (T)	12.20 (2.98)		—	.60	−.10	.43
Social comparison (S—O)	1.61 (4.41)			—	.47	.58
Social reality (S—T)	.20 (2.98)				—	.29
Direct BTAE (K)	6.39 (2.24)					—

Note. Measures and total scores are on a scale of 0–20. The direct measure of the BTAE is on a scale of 0–10. For all variables, *N* = 189. If $r \geq .14$, $p < .05$; if $r \geq .18$, $p < .01$.

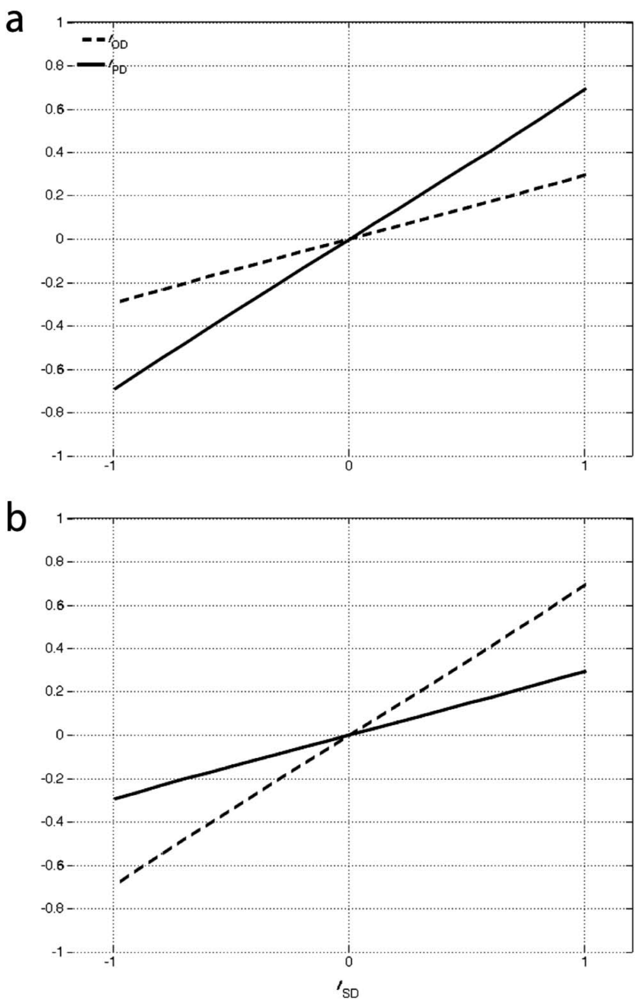

Figure 2. Simulations varying r_{SD} with projection, r_{SO}, and validity, r_{SmeanS}, set at .3 and .7 (Panel a) or .7 and .3 (Panel b).

results (Figure 2, Panel b) show self-effacement error ($r_{\text{OD}} > r_{\text{PD}}$) for positive self-images.

Discussion

Study 1 showed that conventional self-enhancement in personality judgments can be decomposed into a normative and an erroneous component, whether S is inferred from O or vice versa. Although our empirical findings confirm the robustness of self-enhancement bias, they also show that conventional measures overestimate the prevalence of error because they fail to estimate and remove the normative component. How much error is observed critically depends on how much respondents project to others and how typical their own self-judgments are to the aggregates in the group. The outcome is intriguing: while one might think that those individuals who are most prototypical of the group are least likely to self-enhance, the opposite is the case.

Study 2: Performance

Turning from judgments of personality to perceptions of performance, we treat the perceiver's challenge as a decision problem. Estimating their own and the average other's score on a performance test, respondents can predict a favorable or an unfavorable social comparison, knowing that their predictions can be checked for accuracy. By combining self-estimates with other-estimates, true individual performance scores and average performance scores, a decision-theoretic framework allows us to differentiate between those who mistakenly believe to be above average (False Alarms) from those who do so correctly (Hits). Only the former commit a self-enhancement *error*. Likewise, the decision framework permits a distinction between correct and mistaken self-effacers. In Study 2, we estimate the prevalence of these four types with a trivia test, and we test the prediction that conventional measures of self-enhancement are vulnerable to the reverse-inference fallacy. We use correlational analyses and computer simulations to further study the properties of the False Alarm index of self-enhancement error.

Method

Participants. Participants (N = 201) were recruited using Amazon Mechanical Turk. Ten participants, who failed to complete the task or admitted to using outside help, were excluded from analysis. Two additional cases were excluded for reporting more than a single value in response to any of the performance assessment questions. The data of 189 participants remained for analysis (53 female, median age = 25).[4]

Procedure. We prepared an online survey using Qualtrics (2013) and we limited eligibility for participation to residents of the United States. The survey comprised a 20-item sports quiz with 10 items of medium difficulty, 5 easy items, and 5 difficult items, which we adapted from Moore and Small (2007). Interitem reliability was moderate to high (α = .75). After completing the quiz, participants were asked to provide performance estimates for themselves, S ("How many questions do you estimate you answered correctly?") and the average other person, O ("How many questions do you estimate the average other person answered correctly?") in counterbalanced order. Participants' total scores, T, were tallied after data collection as an index of their performance.

To provide an additional measure of the conventional BTAE, participants rated the statement "I am more knowledgeable than the average person" on a scale ranging from 0 (*strongly disagree*) to 10 (*strongly agree*). We refer to this variable as K for *Knowledge*. The trivia task and the question for K were presented in counterbalanced order. We created one random order for the quiz items and used it throughout. Finally, participants reported their gender and age. We queried them about whether they used any materials outside of their own knowledge to answer the questions (Goodman, Cryder, & Cheema, 2013). Participants were then debriefed and given a confirmation code to enter into the survey client to receive compensation for completion.

Results

Conventional analyses. Table 7 shows descriptive statistics for the three primary variables, S, O, and T, the two simple

[4] We conducted two studies of this type, and report the results of one. In the unreported study, the findings and the statistical tests were highly similar, but the trivia task was less reliable.

difference scores of self-enhancement, S—O, S—T, the BTAE index K, as well as all intercorrelations. S judgments were higher than O judgments, $t(188) = 5.02$, $p < .01$, $d = .47$, indicating favorable self-other comparisons. There was, however, no evidence for a social-reality enhancement effect, as S judgments did not differ statistically from T scores, $t(188) < 1$. This lack of a mean-level effect is not an unusual result (Colvin et al., 1995; John & Robins, 1994). Variable K also produced a BTAE; its mean lay above the midpoint of the scale, $t(188) = 5.47$, $p < .01$, $d = .40$.

Most intercorrelations among the primary and the derived measures were positive, as one should expect on empirical and logical grounds; hence we only note exceptions. O judgments were unrelated to test scores; neither high scorers nor low scorers were particularly biased against the average other. Furthermore, the correlation between T and the social reality index S—T was negative, as one should expected on mathematical grounds (Krueger & Wright, 2011; McNemar, 1969), but the effect was surprisingly small.

The data fell into the pattern characteristic of Moore and Small's (2007) differential-regression model (see also Fiedler & Krueger, 2012). Figure 3 shows that S judgments were slightly regressive with respect to T, $b = .93$. This finding attests to the overall accuracy of self-perception. In contrast, O judgments were almost perfectly regressive, $b = .04$. Taken together, these two regressions reveal an important and lawful divergence between the social comparison and the social reality approach. As Moore and Small pointed out, low scorers (low T) are likely to overestimate their own performance (social reality), while believing that they scored worse than others (social comparison), whereas high scorers show the reverse pattern. Finally, S judgments predicted O judgments, $b = .28$, $p < .01$, suggesting social projection. The modesty of this effect leaves ample room for the BTAE to emerge.

Decision-theoretic categories. Using a standard decision-theoretic classification scheme, we sorted respondents as follows. Those who thought they scored better than average, $S > O$, and did score higher than average (median), $T > \bar{T}$, were Hits, H ($N = 79$). Those who thought they scored better than average, but did not, $S > O$ and $T \leq \bar{T}$, were False Alarms, FA, or true self-enhancers ($N = 44$). Those who thought they did worse than average and did, $S \leq O$ and $T \leq \bar{T}$, were Correct Rejections, CR ($N = 52$), and those who thought they did worse than average, but did not, $S \leq O$ and $T > \bar{T}$, were Misses, M, or true self-effacers ($N = 14$). Figure 4 shows the resulting matrix. The correlation between perception and reality was $\Phi = .41$, suggesting a fair degree of accuracy. Notice that the aggregate of the H and FA reflects the self-enhancement bias as seen from social-comparison perspective, whereas only the FA may be said to have committed an error of judgment.

Self-enhancement error. To see how well the conventional indices of self-enhancement predict error, we examined the correlations between their respective difference score measures and a dummy variable, in which a FA (i.e., $S > O \cap T \leq \bar{T}$) was scored as 1 and all other types as 0. The correlations were $r(187) = .22$, .41, and .52, respectively, for the social comparison measure S—O, the social reality measure S—T, and the hybrid measure S—O—T (all $p < .05$).

We then considered the hypothesis that conventional difference-score indices of self-enhancement overdiagnose error.[5] For the social comparison measure, it is easy to see that overdiagnosis is inevitable. If $S > O$, the person is a conventional self-enhancer, but could be either a False Alarm or a Hit in the decision-analytic context. The same is true for the hybrid index. For any False Alarm, the hybrid score will be positive ($S—O—T + \bar{T} > 0$) because $S > O$ and because T, which is subtracted, is smaller than $\bar{T}$, which is added. The reverse is not true. The hybrid score is also positive for a Hit if $S—O > T—\bar{T}$. For the social reality measure, overdiagnosis cannot be shown deductively. A False Alarm ($S > O \cap T \leq \bar{T}$) or a Hit ($S > O \cap T > \bar{T}$) can occur if $S < T$ or if $S > T$.

In our sample, all three conventional measures overdiagnosed self-enhancement. For the social comparison measure, the probability of $S—O > 0$ given a FA was 1, whereas the inverse conditional probability was .36. For the social reality index, the probability of $S—T > 0$ given FA was .77, whereas the inverse conditional probability was .38. For the hybrid index, the probability of $S—O—T + \bar{T} > 0$ given a FA was 1, whereas the inverse conditional probability was .37. Table 7 shows the frequency tables from which these probabilities were calculated.

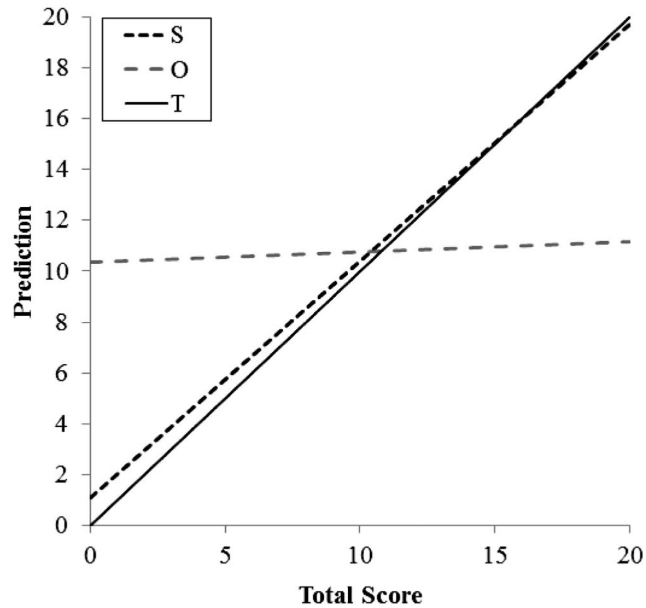

Figure 3. Predicted values generated by regressing self-judgments S, and judgments of the average person O, on total score T (Study 2).

Simulating Categorization

Having opened an empirical window to see how self-judgments, other judgments, and performance data can be combined to identify four types of respondent within a general decision-theoretic framework, we found that the data not only recovered familiar patterns (e.g., Moore's differential regression), but also revealed the hypothesized confound between self-enhancement and inaccuracy in conventional measures. These measures are vulnerable to

[5] This interest is a theoretical choice. Arguably, the misdiagnosis of correct rejectors as biased self-effacers is equally important.

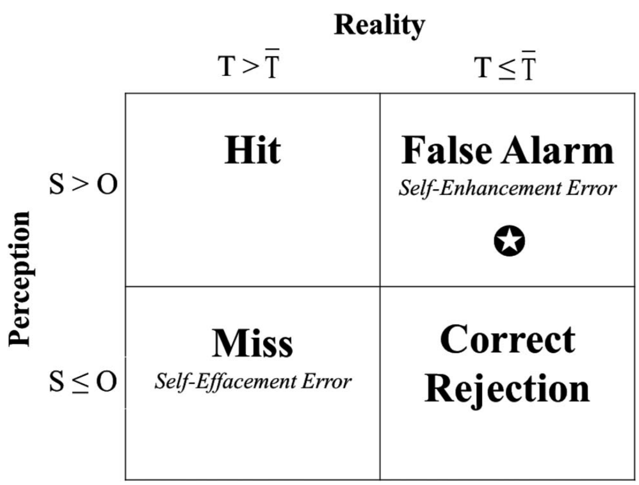

Figure 4. Decision-theoretic categories of self-perception and performance. S = self-judgment; O = other judgment; T = total.

systematically overdiagnosing self-enhancement. For all three conventional measures, we found that a person supposedly showing self-enhancement was more likely to be accurate than inaccurate.

As we did after Study 1, we sought to situate our empirical findings within a broader quantitative context. Therefore, we conducted computer simulations to see how the False Alarm measure compares with the hybrid measure of self-enhancement. Recall that the hybrid measure also represents an effort to integrate the social-comparison approach with the social-reality approach, but that we have expressed concern about the measure's ability to do so.

We performed 11 simulations with 20,000 individuals each. In each simulation, the accuracy correlation between self-judgments S and total score T ranged from 0 to .99 in steps of .1. All other input parameters remained constant and in alignment with the results of our empirical study and past findings. Means (*SD*s) were as follows: S = 14.0 (4.0), O = 12.0 (2.75), and T = 10.0 (3.0).[6] To recreate the relationships among variables found in Study 2, the correlation between S and O judgments (social projection) was set to .2 and the correlation between O judgments and T scores was set to 0.

Panel a of Figure 5 shows the overall effect of judgmental accuracy, r_{ST}, on the relative proportions of the four decision-theoretic categories without regard to the hybrid measure. Not surprisingly, H and CR become more prevalent as accuracy increases, whereas FA and M become less prevalent.

Panel b shows the classifications of respondents with positive hybrid scores ($S—O—T + \bar{T}$). There are now more FA as one would expect. When accuracy is low, the hybrid measure performs best as it is more likely associated with FA than with H. As accuracy increases, however, a positive hybrid score begins to track H rather than FA. In other words, judgmental accuracy moderates the degree to which the hybrid measure correctly detects self-enhancement error. Recall from the discussion of reverse inference that a positive hybrid index score is necessary for a False Alarm. The simulations show that even when there is no correlation between self-judgments and test scores, that is, even in the case in which the hybrid score performs best, 35% of the simulated individuals are correct in their comparatively favorable self-assessments.

Panel c shows the results for negative hybrid scores. Only few individuals commit a self-effacement error (M). Even when accuracy is zero, many of the individuals with a negative hybrid score (ca. 22%) are correct in their comparatively unfavorable self-assessment. As accuracy increases, the hybrid score performs increasingly worse, capturing more CRs and fewer Ms.

We also examined the effects of varying social projection, r_{SO}, while holding accuracy positive and constant. We found that differences in social projection do not affect the relative proportions of the four decision-theoretic categories. This is not surprising because categorization depends on the difference S—O observed for individual respondents, whereas the projection correlation reflects the profile similarity of these two variables over respondents. Recall that that in Study 1, projection was negatively associated with self-enhancement error, such that each respondent was characterized by an idiographic projection correlation.

Study 3: Feedback

We sought to replicate the findings of Study 2 while adding a critical experimental intervention. Half the respondents received their actual test scores after they had estimated them, thereby constraining the context of their decision-making. Estimates of O remained the only variable capable of affecting accuracy and bias. We hypothesized that respondents would use their own T score to predict the score of the average person, O; they would no longer need to rely on their estimate S of their performance. In other words, we expected respondents to use projection to construct O, and to do so by using the best information available, namely T. After feedback, a False Alarm occurs if $T > O \cap T \leq \bar{T}$, that is, if an individual with an average or worse-than-average score places the average person below his own (known) performance.

How might feedback change the incidence of self-enhancement *error*? The outcome would depend on the difference between S and T, and the effect of T on O. Among respondents who initially overestimated their performance, feedback would reduce self-enhancement error unless O estimates were to drop to offset the drop from S to T. Conversely, among respondents who initially underestimated their performance, feedback would *increase* error unless O estimates offset the rise from S to T.

We included a traditional difference-score measure of self-enhancement for the dimension of contentment and a measure of self-esteem. We expected that the contentment-related self-enhancement measure would track the difference-score measure, S—O, in the performance domain, but that it would be unrelated to self-enhancement error, FA. We predicted that self-esteem would be related to self-judgments but not to other judgments or test scores. Self-esteem would be positively correlated with the three conventional measures of self-enhancement because each of these difference-scores contains self-judgments (Krueger & Wright, 2011). Revisiting the Taylor and Brown (1988) proposal, we

[6] We ran two additional simulations with the opposite pattern of means ($S < O < T$), and simulations in which S, O, and T are all equal. The results were analytically predictable and consistent with our conclusions.

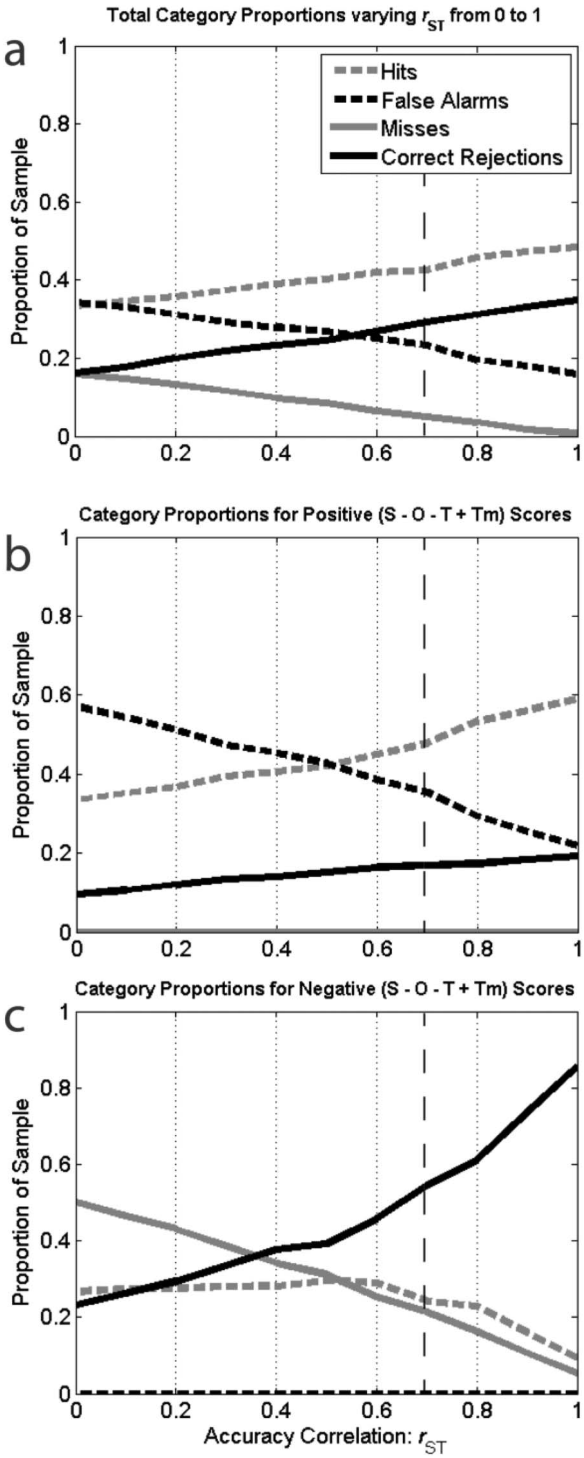

Figure 5. Simulation extending Study 2, where accuracy, r_{ST}, is varied from 0 to 1. The hybrid index is computed as S—O—T + T̄. Panel a represents the proportion of each category in the simulated population. Panel b displays percentages of each category for those with a positive hybrid index score. Panel c displays these proportions for those with a negative hybrid index score. The vertical dashed line represents the simulation resulting from Study 2's empirically observed accuracy correlation ($r_{S,T}$ = .69). S = self-judgment; O = other judgment; T = total.

explored the possibility that self-esteem might be correlated with self-enhancement error.

Method

Participants. Undergraduate participants were recruited at Brown University (*N* = 202). The data of two participants were excluded from analysis because they predicted the average person would only answer 1 out of 30 questions correctly (an extreme outlier). The data of 98 women and 102 men remained for analysis. Exact age information was not collected.

Procedure. All materials were formatted using the Qualtrics survey tool (Qualtrics Research Suite, 2013). Participants first completed the performance and social self-esteem scales of the State Self-Esteem Scale (SSES) (Heatherton & Polivy, 1991). They then completed a 30-item trivia quiz, randomized for each participant, containing 4 easy items, 14 medium-difficulty items, and 12 difficult items (Moore & Small, 2007). Item difficulties were selected to increase score variance while avoiding extreme results. To provide feedback to participants in real time, we used a multiple-choice format with one correct answer and three foils for each question. After completing the quiz, participants were asked to estimate how many questions they answered correctly. Participants in the experimental condition were then told their actual score and asked to estimate how many questions the average test-taker answered correctly. No feedback was given in the control condition: participants simply estimated how many questions the average person answered correctly. Finally, participants were asked to rate how content they and the average person are with their life overall on a Likert scale ranging from *very discontented* (0) to *very contented* (6). We refer to the difference between these two ratings as K for *Kontentment*. The survey concluded with a standard debriefing form.

Results

The general pattern resembled the findings of Study 2 (see Tables 8 and 9). Difference scores were correlated with their individual components and with one another as one should expect on mathematical grounds. In the control condition, the pattern of differential regression emerged again, such that S judgments tracked T scores more closely than did O judgments. The K variable showed statistically significant self-enhancement both in

Table 8
Descriptive Statistics and Intercorrelations (Study 3 Control Condition)

Measures	*M* (*SD*)	Correlation O	T	S—O	S—T	K
Self-judgment (S)	17.03 (5.18)	.63	.27	.51	.85	.06
Other-judgment (O)	16.79 (4.78)	—	.11	−.35	.57	−.19
Total score (T)	21.0 (3.41)		—	.21	−.28	.00
Social comparison (S—O)	.24 (4.07)			—	.39	.29
Social reality (S—T)	−3.97 (4.92)					.07
Direct BTAE (K)	.49 (1.11)					—

Note. Measures and total scores are on a scale of 0–30. The direct measure of the BTAE is on a scale of 0–6. For all variables, *N* = 100. If $r \geq .20$, $p < .05$; if $r \geq .25$, $p < .01$.

Table 9
Descriptive Statistics and Intercorrelations (Study 3 Feedback Condition)

Measures	*M* (*SD*)	Correlation				
		O	T	S—O	S—T	K
Self-judgment (S)	17.06 (5.17)	.13	.54	.82	.80	.00
Other-judgment (O)	18.90 (3.63)	—	.29	−.45	−.05	−.03
Total score (T)	21.15 (3.59)		—	.32	−.06	.00
Social comparison (S—O)	−1.80 (5.46)			—	.75	.02
Social reality (S—T)	−4.09 (4.14)				—	.00
Direct BTAE (K)	.64 (1.10)					—

Note. Measures and total scores are on a scale of 0–30. The direct measure of the BTAE is on a scale of 0–6. For all variables, $N = 100$. If $r \geq .20$, $p < .05$; if $r \geq .25$, $p < .01$.

the control, $t(99) = 4.43$, $p < .01$, $d = .52$, and the feedback condition, $t(99) = 5.84$, $p < .01$, $d = .67$.

Effects of feedback. Tables 8 and 9 show the means (*SD*s) for the primary variables and the difference scores, as well as their intercorrelations, respectively, for the control (no-feedback) and the feedback condition. Our projection hypothesis stated that in the absence of feedback, respondents would base their O judgments on their own S judgments. Consequently, the correlation between O and S should be greater than the correlation between O and T. This was the case, $r_{SO} = .63, > r_{OT} = .11$, $Z = 4.88$, $p < .01$ (Steiger, 1980). Conversely, when information regarding test scores was fed back, projection implies that the correlation between O and T would be greater than the correlation between O and S. This part of the projection hypothesis had modest support, $r_{SO} = .13, < r_{TO} = .29$, $Z = 1.69$, $p < .09$.

Consistent with the projection hypothesis was the finding that O judgments were higher in the feedback condition ($M = 18.86$) than in the control condition, ($M = 16.79$), $t(198) = 3.77$, $p < 01$, whereas the means of S and T did not differ between conditions, $t < 1$. This mean-level effect in O reflected the predicted shift from the use of S judgments to the use of *T* scores as the preferred basis for estimates of O. In both conditions, S judgments were significantly lower than *T* scores (control: $t(99) = 8.10$, $p < .01$; feedback: $t(99) = 9.89$, $p < .01$). The difference between S and T in the feedback condition was nominally greater than the difference in O judgments between conditions. This constellation suggested an *increase* in self-enhancement error after feedback. The logic is this: When respondents had underestimated their performance (S < T), feedback raises the number of individuals who think they scored higher than average. This change increases both Hits and False Alarms unless O estimates increase as much as S estimates do. Because projection to others is not perfect, we can only expect O estimates to rise moderately after feedback regarding S. This is, indeed, what we found.

In the control condition, the proportions for the four classifications were 20% H, 26% FA, 21% M, and 33% CR. In the feedback condition, the proportions were 42% H, 35% FA, 6% M, and 17% CR. Feedback thus boosted both Hits and False Alarms. Because the increase in H was greater than the increase in FA, there was a modest increase in overall accuracy. Whereas in the control condition, the proportion of correct classifications (H + CR) was 53%, ($\Phi = .05$), the corresponding proportion after feedback was 59% ($\Phi = .24$), $p < .19$.

These findings show that the decision-theoretic framework is not only suited to separate bias from accuracy, but is also sensitive to changes in both bias and accuracy after the introduction of more information. Of course, the findings of this study can only provide an empirical snapshot. As we had done after the first two studies, we explored the regularities to be expected in this framework with simulations. First, however, we revisit the issue of how conventional indices of bias invite a reverse-inference problem.

Diagnosis of error. In the control condition, all three conventional measures were correlated with FA (vs. the other three categories), $r(98) = .41$, .34, and .57, respectively, for the social comparison measure S—O, the social reality measure S—T, and the hybrid measure S—O—T (all $p < .05$). The probability of a positive social comparison or a positive hybrid score given a FA was again 1; thus, making overdiagnosis and the reverse inference fallacy virtually inevitable. Given a FA, the probability of S > O was .58 and the probability of a positive hybrid score was .57. There was no overdiagnosis for the social reality measure. The probability of S > T given a FA was .46, whereas the inverse probability (FA given S > T) was .60.

In the feedback condition, T scores replaced S judgments so that social comparison reduced to T—O and the hybrid measure reduced to $\bar{T} - O$. The social reality measure reduced to T—T, or a constant of 0. Self-enhancement error was given by a False Alarm = $T > O \cap T \leq \bar{T}$. Over respondents, False Alarms (vs. all other categories) were not correlated with T—O, $r(98) = .12$, but they were correlated with the hybrid measure, $r(98) = .49$. Given a FA, the probability of a positive conventional index, T—O or $\bar{T} -$ O, was 1. The inverse conditional probability, however, was merely .46 in either case. In short, the data reveal a massive reverse inference problem. Individuals charged with self-enhancement by conventional measures were more likely accurate in their self-perception than self-aggrandizing.

Self-esteem. We hypothesized that self-esteem would be associated with measures of self-enhancement. The SSES yielded separate scores for performance and social self-esteem. We averaged these scores because they were sufficiently correlated (Control: $r = .89$; Feedback: $r = .62$). As predicted, self-esteem was correlated with S, $r = .31$, $p < .01$, but not with O, $r = -.01$ or T, $r = .04$, in the control condition. As a consequence, self-esteem predicted all difference-score indices of self-enhancement: $r = .38$ (S—O: social comparison), .29 (S—T: social reality), and .33 (hybrid). Self-esteem also predicted self-enhancement error, FA, $r = .26$ (all $p < .01$, all $df = 98$) but not correct enhancement, H, $r = .18$. The correlation between self-esteem and self-enhancement error remained when correlations involving H were controlled, $r = .33$. The feedback condition yielded a similar pattern. Self-esteem was correlated with S, $r = .40$, but not with O, $r = -.10$ or T, $r = .01$. However, self-esteem predicted the modified social comparison measure T—O, $r = .42$, and it was marginally correlated with FA, $r = .16$, $p = .06$, and $r = .19$, $p <$ .19, when correlations with H were controlled. There was no correlation with correct enhancement, H, $r = -.01$. In both conditions, self-esteem predicted the difference score variable K (*Kontentment*), $r = .40$ and .20 in the control and the feedback condition, respectively (both $p < .05$, all $df = 98$). In short, the correlations involving the measure of self-esteem were largely as predicted. The most intriguing result was the positive partial correlation with self-enhancement error, a finding that is consistent with the Taylor-and-Brown hypothesis.

Simulating postfeedback self-enhancement. A key objective of Study 3 was to show that lacking precise information about their own performance and the performance of the average other, respondents construct O judgments projectively on the basis of their self-judgments. Once available, however, the true value T is a better anchor on which to base a prediction for the average person. This is so because one's own performance is part of the average performance. Our sample data supported this idea and they also showed increases in both self-enhancement error and categorization accuracy. As empirical snapshots, these findings are contingent on the vagaries of sampling respondents and tasks. Again, we used simulations to ask how the relative differences between S and T combine with changes in O would affect accuracy and bias over a variety of settings.

For two scenarios, we performed 11 simulations with 10,000 individuals each. All samples were integers between 0 and 30. Correlations between simulated variables were set to empirically plausible values: .5, .2, and 0 for r_{ST}, r_{SO}, and r_{OT}, respectively. Because judgments of the average other were the target variable in Study 3, we varied the mean value of O from 10 to 20 in steps of one (*SD* = 4.0). The panels of Figure 6 show the variation in O on the x-axes. The figure displays category proportions and accuracy. Accuracy coefficients Φ were computed for each setting of O and are plotted as a separate line in all panels. The simulations shown in Panel a model the empirically observed case in which self-judgments (*M* = 11, *SD* = 5.0) are lower than true scores (*M* = 15, *SD* = 2.8). Panel b shows the results when the simulated individuals are categorized using T—O instead of S—O. Finally, the simulations shown in Panel c, model the idealized case in which self-judgments (*M* = 19, *SD* = 5.0) are greater than true scores and individuals are categorized using S—O.

Classification using T—O (Panel b) results in greater accuracy than classification using S—O (Panels a and c), regardless of the relation between S and T. This result is particularly interesting because we empirically observed a higher incidence of self-enhancement error (FA) when feedback T was present instead of absent. Our simulations replicate this effect. Unless O is very high, the proportions of the simulated samples categorized as FA are higher in Panel a than in Panel b. However, when S > T (Panel c), the proportions of FA are higher for the simulated control condition, S—O (Panel c), than for the feedback condition (Panel b).[7]

A simultaneous increase in both accuracy, Φ, and self-enhancement error, FA, is possible, but it may seem counterintuitive because it is empirically rare. We now show how feedback—and thereby replacing S with T—yields more Hits and more False Alarms. Recall that in our sample, T > O > S. Feedback replaces S with T, and therefore (T—O) > (S—O) in most cases. When S < T, providing feedback raises the decision criterion and thereby reduces bias (see Figure 4). Feedback turns some Misses into Hits, and some Correct Rejections into False Alarms. Because the reality of performance (T relative to $\bar{T}$) does not change, movement in the table can only be vertical. Recall, for example, that a CR requires $S < O \cap T \leq \bar{T}$. Because T > O in most cases, some individuals will be recategorized from S < O to T > O. $T \leq \bar{T}$ does not change, however; this restricts the only possible recategorization of a CR to a FA.

Now we can consider our empirical finding that the increase in H is greater than the increase in FA (resulting in a more positive Φ). Before feedback, a Hit requires $S > O \cap T > \bar{T}$. After feedback, T replaces S: a Hit now requires $T > O \cap T > \bar{T}$. Since the inequality $T > \bar{T}$ dictates that T is relatively large, we can infer that T is likely larger than another random variable, in this case O. By contrast, a False Alarm following feedback requires $T > O \cap T \leq \bar{T}$. Here, T tends to be small, and therefore, less likely to be larger than some other random variable O. As such, replacing S with T when recategorizing individuals *must* result in a larger recategorization of Misses to Hits than Correct Rejections to False Alarms.

We can now ask whether respondents can manage the accuracy of their predictions by strategically adopting a larger or a smaller self-enhancement bias by changing their O judgments. The simulations show that it depends on whether respondents underestimate (S < T) or overestimate (S > T) their own performance. The outcomes shown in Panel a (where S < T) point to increases in accuracy with lowered O; the outcomes shown in Panel c (S > T) are the reverse. In other words, social comparisons can be used strategically to compensate for social reality. Underestimators can gain accuracy by being more liberal in their social comparisons; overestimators gain accuracy by being stricter.

General Discussion

Conventional approaches to the study of self-enhancement conflate judgmental bias with error and inaccuracy. They leave little room for the perceiver to emerge as a rational information processor. We present a theoretical rationale and two methods for the separation of rational (or accurate) judgments from erroneous ones. We tested the two novel measurement indices in three studies and three sets of simulations. Our findings show new opportunities for fertile research on self-enhancement.

The Projection Index

In the domain of personality judgment, we used the logic of social projection and statistical regression to estimate how a rational, though not omniscient, person might construct judgments of the average other person in an uncertain environment. We predicted and found that such a person would likely self-enhance. By subtracting the rationally justified component of self-enhancement from the overall effect, we isolated self-enhancement *error*, that is, enhancement beyond the rationally justified. Self-enhancement error varies in systematic ways. It is common among people with a positive self-image who project this image less strongly to the average person than is warranted by the cue validity of their self-judgments. The projection-based index achieves the separation of bias and error without a reality criterion of what the perceiver "is really like," but it is anchored in social reality in that it considers the actual similarity between the perceiver's self-judgments and the aggregate self-judgments in the group.

[7] Using higher correlations among S, O, and T (.5) yields a decrease in FA and Φ when categorizing using T-O and a slight increase when using S-O. Using intercorrelations of 0 causes Φ to drop to zero when using S-O.

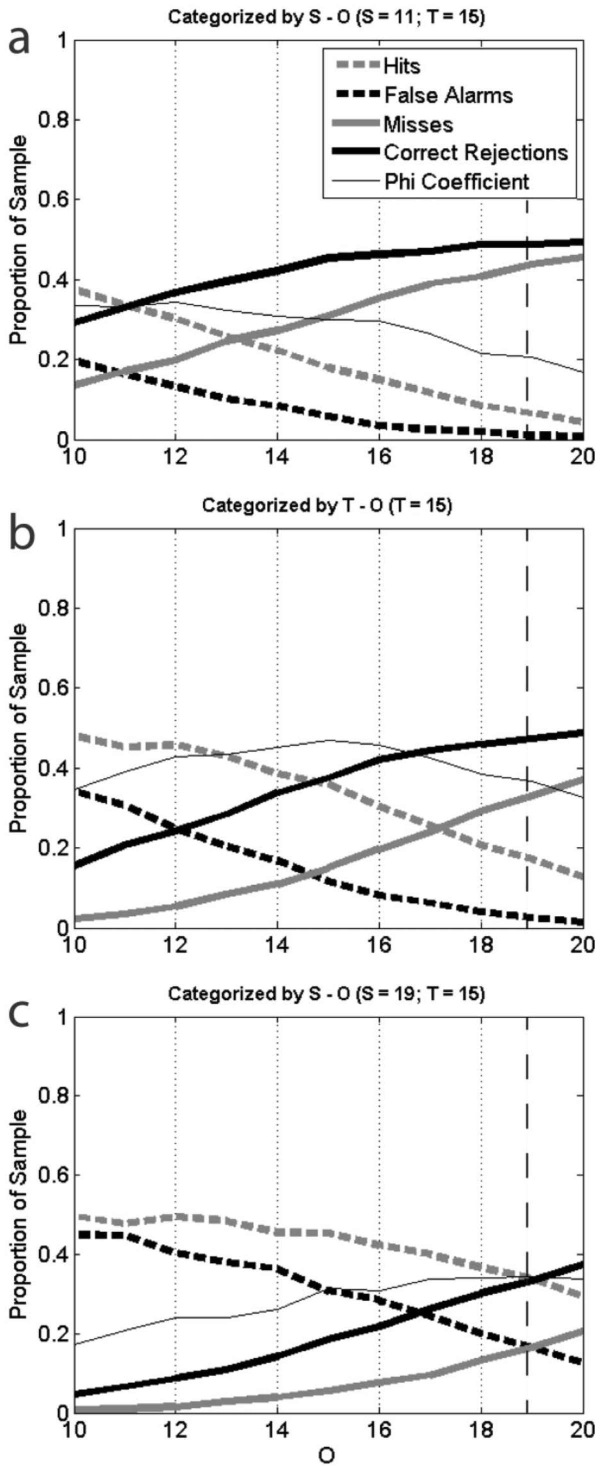

Figure 6. Simulation extending Study 3, where rating of the average person, O, is varied from 10 to 20. Each panel displays category proportions and accuracy, Φ. Panel a displays the empirically observed scenario, where S < T. Panel b displays category proportions and accuracy when categorized using T—O. Panel c displays the idealized scenario, where S > T. The vertical dashed line represents the simulation resulting from Study 3's empirically observed O (18.9). S = self-judgment; O = other judgment; T = total.

The Decision-Theoretic Index

For domains of performance, we developed a decision-theoretic method, which uses test scores as an objective reality criterion. Respondents who falsely believe to be better than average commit a Type I error, whereas those who correctly believe to be better than average are justifiably biased. Both indices focus on the individual perceiver's data, thereby guarding against the ecological fallacy; they recognize that what is found at the group level may not characterize all individuals (Robinson, 1950).

We explored the properties of the decision-theoretic measure further by conducting an additional study with university students as respondents and their college peers as the target group for social predictions (Heck, Krueger, & Sachs, unpublished). By constraining the sample of respondents and the population from which they were drawn, we hoped to corroborate the validity of our approach. Making an additional change, we told some participants that the average student had scored either very high or quite low on the test. That way, any strategic construction of self-enhancement could occur at the level of self-judgments. The overall pattern of results was very similar to that seen in Study 3. Intriguingly, we also observed that the incidence of False Alarms relative to Hits fell as the alleged performance of others rose. We conjecture that telling respondents that the average other performed well raises the perceived costliness of making an error. Respondents might worry about being perceived as hubristic if they claimed to have received an even higher score. This brief sketch of the follow-up study suggests that the decision-theoretic approach can be used to (a) explore the unique contributions of self-judgments and other judgments to self-enhancement, and to (b) explore the role of motivated reasoning.

Studies 2 and 3 provided an additional opportunity to examine the relationship between bias and error in a way suggested by Epley and Dunning (2006). For each of the four decision-theoretic types, we computed accuracy correlations over respondents. In Study 2, we found that among the self-enhancers, the self-judgments of respondents who were categorized as Hits were more calibrated, r_{ST} = .80, than the self-judgments of those who were categorized as False Alarms, r_{ST} = .095. Among the self-effacers, accuracy correlations were intermediate in size, r_{ST} = .38 and .58, respectively, for Misses and Correct Rejections. Study 3 yielded the same pattern. While corroborating Kruger and Dunning's (1999) hypothesis that low-scoring individuals (FA and CR) are less calibrated than high scorers, we see that it is those who commit self-enhancement errors (FA) that are most lacking in calibration.

Epley and Dunning (2006) found that when predicting their own likelihood to act prosocially (e.g., by voting), self-judgments were sensitive to "case-based" information (e.g., one's past behavior) instead of distributional information (i.e., base rates). Judgments of others showed the opposite asymmetry, and were more accurate. This work is among the most prominent examples for the view that self-judgments are overall more biased than other judgments. As noted above, our projection-based measure does not speak to this potential difference. In the decision-theoretic framework, we find in both studies that self-judgments were more accurate than other judgments (that underestimated average performance), while calibration could only be studied for S but not O (because the latter did not vary over respondents).

Implications for Rationality

Since Taylor and Brown's (1988) influential review, self-enhancement has been widely considered a footprint of failed social reasoning. Taylor and Brown suggested that along with the illusions of control and personal invulnerability, self-enhancement reveals deeply rooted design flaws of the social mind. When, as is generally the case in the social comparison literature, attention is limited to the group-level effect, the door is open to an ecological fallacy. Self-enhancement then becomes a property of the abstract social perceiver; it becomes "something that people do." Indeed, Svenson (1981) entitled his foundational article with the rhetorical question "Are we all less risky and more skillful than our fellow drivers?"

Measures that separate error from bias protect researchers from making blanket claims about "what people do" and hence from the ecological fallacy. We must return, however, to the question of what these new measures say about the (ir)rationality of the mistaken self-enhancers and social judges in general. The projection-based measure describes each respondent in terms of the component mix of bias and error.

One approach to rationality turns on the concept of correspondence (Hastie & Rasinski, 1988). Correspondence refers to judgmental accuracy, which may be either a match between judgments and external criteria or between judgments and other rationally obtained estimates (Jussim, Stevens, & Salib, 2012). The projection-based measure may be used to ask how rational bias and error are related to respondents' accuracy when judging the average person. In Study 1, we find that the magnitude of rational bias, $b_{SD} - b_{PD}$, is unrelated to accuracy (i.e., the correlation between O judgments and mean S judgments in the sample) over individuals, $r(79) = .06$, whereas the size of the self-enhancement error, $b_{PD} - b_{OD}$, is negatively correlated with accuracy, $r(79) = -.74$. This result supports our claim that self-enhancement error signals a lack of accuracy in social perception, which is consistent with the first part of Taylor and Brown's (1988) hypothesis.[8]

The second, and more provocative, part of the Taylor-and-Brown hypothesis is that self-enhancement is adaptive. Whereas Taylor and Brown presented only evidence obtained with the response-bias measure of self-enhancement (S—O), we found that errors predicted self-esteem. While this might be the most direct piece of evidence for the Taylor-and-Brown hypothesis to date, we hasten to point out that measures of self-esteem and self-enhancement are both obtained from the same individuals (Krueger & Wright, 2011). It will be interesting to see whether self-enhancement error predicts measures of adaptivenss obtained from independent sources (e.g., social acceptance).

An alternative definition of rationality refers to the coherence of judgments. Individuals are being rational if their own judgments do not contradict one another or if it can be shown that they will not contradict one another if these individuals keep making more judgments in the same way (Dawes, 1988). By this definition, neither the conventional nor the new projection measure of self-enhancement reveals irrationality.

The decision-theoretic measure also distinguishes bias and error. Judgmental (or response) bias is the threshold individuals set to call a positive result. This threshold is given by the relation between two estimates, S and O. Respondents who think that S > O bet that their judgments will result in a Hit and not a False Alarm; they are ruling out being a Miss or a Correct Rejection. The conventional social comparison measure S—O captures this kind of bias. The new measure asks more concretely whether the individual's claim that S > O is false. The identification of such a false positive error requires a truth criterion.

A person who shows self-enhancement error is being inaccurate, but not necessarily irrational (Einhorn, 1986). Instead, this person attempts to manage two types of error: Type I errors (False Alarms) and Type II errors (Misses). This person will be sensitive to the benefits (utility) of correct decisions and the costs (disutility) of wrong ones (Haselton, Bryant et al., 2009; Lynn & Feldman Barrett, 2014). A person may set a liberal threshold for claiming to be better than average inasmuch as a Miss seems more damaging than a False Alarm. In the current cultural climate, many individuals may think there is little harm in laying a claim to relative superiority. Given that many people strive to be better than average, they might as well predict that they have reached their goal while still being in a state of uncertainty. Doing the opposite would seem odd. Suppose Jerry intends to produce superior ice cream and he knows that his prediction of whether he actually does produce better ice cream than Ben does will be checked against reality. For him to say "I want to make a superior product but I think I make an inferior one" sounds strange and self-defeating if the truth is still uncertain. Likewise, there is little to be gained from accurately predicting one's own inferiority if one's goal is to be superior. From this perspective, the setting of the threshold S—O should be highly sensitive to a judge's motivation to excel.

Once the truth is revealed, and Jerry turns out to be in error, he need not be called irrational. He made a judgment call under uncertainty, using the information he had and consulting his predictions of how he would feel about the possible results. If Jerry were not granted the privilege of making a mistake under uncertainty, any nonomniscient decision-maker would be judged in the light of outcomes they were unable to foresee. However, such judgments would amount to a crude form of outcome bias (Baron & Hershey, 1988; Krueger & Acevedo, 2007), and outcome bias is a stock example of incoherence irrationality (Dawes, 1988).

The Broader Judgment-and-Decision-Making Context

Our theoretical aim to distinguish between bias and inaccuracy (error) continues a longstanding topic in the psychology of judgment and decision-making. Some of the signature phenomena of human irrationality have been criticized for being too liberal in attributing fallacious thinking to research participants (cf. Krueger & Funder, 2004). Research on social projection treated any perceived similarity between the self and others as a "false consensus effect" (Ross, Greene, & House, 1977) until Hoch (1987) and Dawes, Faust, and Meehl (1989) showed empirically and analytically that most people are indeed similar to one another. Their insights informed the development of our projection-based index

[8] There is no reason to think that rational bias must be positively correlated with accuracy. The same regression weight was used for all respondents when estimating the rational judgment of the average other. The correlation between (b_{SD} - b_{PD}) and accuracy is, therefore, sensitive to variation in S, and this variable, by itself, says little about accuracy. While rationality may be reasonably correlated with accuracy, the magnitude of rationally defensible bias need not be.

of self-enhancement. More recently, Harris and Hahn (2011) argued that standard indices of the so-called illusion of own invulnerability overdiagnose human error. To the extent that the catastrophic events of interest (e.g., maiming accidents or debilitating diseases) are rare, those individuals who have not experienced them should indeed feel that they are less vulnerable than the average person. Finally, Gino, Sharek, and Moore (2011) contextualized the illusion of control and showed that the standard illusion emerges only when there is no actual control at all. As actual control increases, overestimation errors diminish and ultimately turn into underestimation, just as the law of regression demands. Placing the study of self-enhancement in this tradition, we recognize the importance of studying human error, while calling for the development and use of more sensitive measures. We hope that researchers, regardless of their theoretical commitments, will find it advantageous to use such measures.

Caveats and Conclusions

Our proposed measures may not be suitable for the identification of latent population parameters. There is great variation in the possible composition of the samples of respondents or tasks. Indeed, our finding in Study 3 that most respondents underestimated performance may be atypical. Nonetheless, we found that even here many individuals could be identified as mistaken self-enhancers.

Empirical results of any given study can only provide a snapshot of what may happen. Therefore, we conducted computer simulations to show what kinds or patterns are possible. Simulations are indispensable for the demonstration of dependencies and constraints built into the ecological and mathematical structure of the problem. They provide the background for the figural signal of the empirical data. We chose not to constrain the magnitude and direction of our parameters of interest; instead, the parameters of our simulations cast a wide net over many possible empirical scenarios. By showing what might or could be the case, simulations help overcome the single-mindedness of comparing empirical observations with null models and nothing else. When theoretical predictions can be made about the effect of context on self- and social perception, simulations can model and quantify these predictions. When empirical data on context effects are available, one may ask whether they may be identified in the terrain described by the simulations.

In closing, we reiterate that our proposal takes no position on longstanding debates over the processes or moderators of self-enhancement. Instead, our measures offer researchers tools to pursue theoretical and pragmatic questions such as: Is self-enhancement adaptive? Does it depend on the cultural environment or enduring characteristics of the person? What are the cognitive and motivational processes that regulate how individuals set their own parameters of decision-making, and how responsive are individuals to incentives to maximize accuracy or positive impressions?

References

Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology, 49,* 1621–1630. http://dx.doi.org/10.1037/0022-3514.49.6.1621

Alicke, M. D., & Sedikides, C. (Eds.). (2011). *Handbook of self-enhancement and self-protection*. New York, NY: Guilford Press.

Allison, S. T., Messick, D. M., & Goethals, G. R. (1989). On being better but not smarter than others: The Muhammad Ali effect. *Social Cognition, 7,* 275–295. http://dx.doi.org/10.1521/soco.1989.7.3.275

Allport, F. H. (1924). *Social psychology*. Boston, MA: Houghton Mifflin.

Ames, D. R. (2004). Strategies for social inference: A similarity contingency model of projection and stereotyping in attribute prevalence estimates. *Journal of Personality and Social Psychology, 87,* 573–585. http://dx.doi.org/10.1037/0022-3514.87.5.573

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York, NY: Macmillan.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica, 44,* 211–233. http://dx.doi.org/10.1016/0001-6918(80)90046-3

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology, 54,* 569–579. http://dx.doi.org/10.1037/0022-3514.54.4.569

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5,* 323–370. http://dx.doi.org/10.1037/1089-2680.5.4.323

Brambilla, M., & Leach, C. (2014). On the importance of being moral: The distinctive role of morality in social judgment. *Social Cognition, 32,* 397–408. http://dx.doi.org/10.1521/soco.2014.32.4.397

Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. *Social Cognition, 4,* 353–376. http://dx.doi.org/10.1521/soco.1986.4.4.353

Brown, J. D. (2010). Across the (not so) great divide: Cultural similarities in self-evaluative processes. *Social and Personality Psychology Compass, 4,* 318–330. http://dx.doi.org/10.1111/j.1751-9004.2010.00267.x

Brown, J. D. (2012). Understanding the better than average effect: Motives (still) matter. *Personality and Social Psychology Bulletin, 38,* 209–219. http://dx.doi.org/10.1177/0146167211432763

Chambers, J. R., & Windschitl, P. D. (2004). Biases in social comparative judgments: The role of nonmotivated factors in above-average and comparative-optimism effects. *Psychological Bulletin, 130,* 813–838.

Cho, J. C., & Knowles, E. D. (2013). I'm like you and you're like me: Social projection and self-stereotyping both help explain self–other correspondence. *Journal of Personality and Social Psychology, 104,* 444–456. http://dx.doi.org/10.1037/a0031017

Colvin, C. R., Block, J., & Funder, D. C. (1995). Overly positive self-evaluations and personality: Negative implications for mental health. *Journal of Personality and Social Psychology, 68,* 1152–1162. http://dx.doi.org/10.1037/0022-3514.68.6.1152

Dawes, R. M. (1988). *Rational choice in an uncertain world*. San Diego, CA: Harcourt, Brace, Jovanovich.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243,* 1668–1674. http://dx.doi.org/10.1126/science.2648573

DiDonato, T. E., Ullrich, J., & Krueger, J. I. (2011). Social perception as induction and inference: An integrative model of intergroup differentiation, ingroup favoritism, and differential accuracy. *Journal of Personality and Social Psychology, 100,* 66–83. http://dx.doi.org/10.1037/a0021051

Dunning, D. (2012). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York, NY: Psychology Press.

Einhorn, H. J. (1986). Accepting error to make less error. *Journal of Personality Assessment, 50,* 387–395. http://dx.doi.org/10.1207/s15327752jpa5003_8

Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments produced by errors in self- or social prediction? *Journal of Personality and Social Psychology, 79,* 861–875. http://dx.doi.org/10.1037/0022-3514.79.6.861

Epley, N., & Dunning, D. (2006). The mixed blessings of self-knowledge in behavioral prediction: Enhanced discrimination but exacerbated bias. *Personality and Social Psychology Bulletin, 32,* 641–655. http://dx.doi.org/10.1177/0146167205284007

Eriksson, K., & Funcke, A. (2014). Humble self-enhancement religiosity and the Better-Than-Average Effect. *Social Psychological & Personality Science, 5,* 76–83. http://dx.doi.org/10.1177/1948550613484179

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7,* 117–140.

Fiedler, K. (1996). Explaining and simulating judgment biases as an aggregation phenomenon in probabilistic, multiple-cue environments. *Psychological Review, 103,* 193–214. http://dx.doi.org/10.1037/0033-295X.103.1.193

Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review, 107,* 659–676. http://dx.doi.org/10.1037/0033-295X.107.4.659

Fiedler, K., & Krueger, J. I. (2012). More than an artifact: Regression as a theoretical construct. In J. I. Krueger (Ed.), *Social judgment and decision-making* (pp. 171–189). New York, NY: Psychology Press.

Fiedler, K., & Unkelbach, C. (2014). Regressive judgment: Implications of a universal property of the empirical world. *Current Directions in Psychological Science, 23,* 361–367. http://dx.doi.org/10.1177/0963721414546330

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102,* 652–670. http://dx.doi.org/10.1037/0033-295X.102.4.652

Galesic, M., Olsson, H., & Rieskamp, J. (2012). Social sampling explains apparent biases in judgments of social environments. *Psychological Science, 23,* 1515–1523. http://dx.doi.org/10.1177/0956797612445313

Gino, F., Sharek, Z., & Moore, D. A. (2011). Keeping the illusion of control under control: Ceilings, floors, and imperfect calibration. *Organizational Behavior and Human Decision Processes, 114,* 104–114. http://dx.doi.org/10.1016/j.obhdp.2010.10.002

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26,* 213–224. http://dx.doi.org/10.1002/bdm.1753

Greenwald, A. G. (2012). There is nothing so theoretical as a good method. *Perspectives on Psychological Science, 7,* 99–108. http://dx.doi.org/10.1177/1745691611434210

Harris, A. J. L., & Hahn, U. (2011). Unrealistic optimism about future life events: A cautionary note. *Psychological Review, 118,* 135–154. http://dx.doi.org/10.1037/a0020997

Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., & Moore, T. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition, 27,* 733–763. http://dx.doi.org/10.1521/soco.2009.27.5.733

Hastie, R., & Rasinski, K. A. (1988). The concept of accuracy in social judgment. In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 193–208). Cambridge, United Kingdom: Cambridge University Press.

Haynes, S. N., Richard, D., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7,* 238–247. http://dx.doi.org/10.1037/1040-3590.7.3.238

Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60,* 895–910. http://dx.doi.org/10.1037/0022-3514.60.6.895

Heck, P. R., Krueger, J. I., & Sachs, J. F. (unpublished). *Decision-theoretic self-enhancement after false performance feedback*. Brown University, Providence, RI.

Heine, S. J., & Hamamura, T. (2007). In search of East Asian self-enhancement. *Personality and Social Psychology Review, 11,* 4–27. http://dx.doi.org/10.1177/1088868306294587

Hoch, S. J. (1987). Perceived consensus and predictive accuracy: The pros and cons of projection. *Journal of Personality and Social Psychology, 53,* 221–234. http://dx.doi.org/10.1037/0022-3514.53.2.221

Hoorens, V., Pandelaere, M., Oldersma, F., & Sedikides, C. (2012). The Hubris hypothesis: You can self-enhance, but you'd better not show it. *Journal of Personality, 80,* 1237–1274. http://dx.doi.org/10.1111/j.1467-6494.2011.00759.x

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. http://dx.doi.org/10.1037/0022-3514.66.1.206

Jussim, L., Stevens, S. T., & Salib, E. R. (2012). The strengths of social judgment: A review based on the goodness of fit index. In J. I. Krueger (Ed.), *Social judgment and decision making* (pp. 97–114). New York, NY: Psychology Press.

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

Klar, Y., & Giladi, E. E. (1999). Are most people happier than their peers, or are they just happy? *Personality and Social Psychology Bulletin, 25,* 586–595. http://dx.doi.org/10.1177/0146167299025005004

Krueger, J. I. (1998a). On the perception of social consensus. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 163–240). San Diego, CA: Academic Press.

Krueger, J. I. (1998b). Enhancement bias in the description of self and others. *Personality and Social Psychology Bulletin, 24,* 505–516. http://dx.doi.org/10.1177/0146167298245006

Krueger, J. I. (2007). From social projection to social behaviour. *European Review of Social Psychology, 18,* 1–35. http://dx.doi.org/10.1080/10463280701284645

Krueger, J. I. (in press). Reverse inference. In S. O. Lilienfeld & I. D. Waldman (Eds.), *Psychological science under scrutiny: Recent challenges and proposed solutions*. New York, NY: Wiley.

Krueger, J. I., & Acevedo, M. (2007). Perceptions of self and other in the prisoner's dilemma: Outcome bias and evidential reasoning. *The American Journal of Psychology, 120,* 593–618.

Krueger, J. I., & Chen, L. J. (2014). The first cut is the deepest: Effects of social projection and dialectical bootstrapping on judgmental accuracy. *Social Cognition, 32,* 315–336. http://dx.doi.org/10.1521/soco.2014.32.4.315

Krueger, J. I., Freestone, D., & MacInnis, M. L. (2013). Comparisons in research and reasoning: Toward an integrative theory of social induction. *New Ideas in Psychology, 31,* 73–86. http://dx.doi.org/10.1016/j.newideapsych.2012.11.002

Krueger, J. I., & Funder, D. C. (2004). Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences, 27,* 313–327. http://dx.doi.org/10.1017/S0140525X04000081

Krueger, J. I., Ham, J. J., & Linford, K. M. (1996). Perceptions of behavioral consistency: Are people aware of the actor-observer effect? *Psychological Science, 7,* 259–264. http://dx.doi.org/10.1111/j.1467-9280.1996.tb00371.x

Krueger, J., & Mueller, R. A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology, 82,* 180–188. http://dx.doi.org/10.1037/0022-3514.82.2.180

Krueger, J. I., & Wright, J. C. (2011). Measurement of self-enhancement (and self-protection). In M. D. Alicke & C. Sedikides (Eds.), *Handbook of self-enhancement and self-protection* (pp. 472–494). New York, NY: Guilford Press.

Kruger, J. (1999). Lake Wobegon be gone! The "below-average effect" and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology, 77,* 221–232. http://dx.doi.org/10.1037/0022-3514.77.2.221

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77,* 1121–1134. http://dx.doi.org/10.1037/0022-3514.77.6.1121

Kruglanski, A. W. (1989). The psychology of being "right": The problem of accuracy in social perception and cognition. *Psychological Bulletin, 106,* 395–409. http://dx.doi.org/10.1037/0033-2909.106.3.395

Kurt, A., & Paulhus, D. L. (2008). Moderators of the adaptiveness of self-enhancement: Operationalization, motivational domain, adjustment facet, and evaluator. *Journal of Research in Personality, 42,* 839–853. http://dx.doi.org/10.1016/j.jrp.2007.11.005

Kwan, V. S. Y., John, O. P., Kenny, D. A., Bond, M. H., & Robins, R. W. (2004). Reconceptualizing individual differences in self-enhancement bias: An interpersonal approach. *Psychological Review, 111,* 94–110. http://dx.doi.org/10.1037/0033-295X.111.1.94

Larrick, R. P., Burson, K. A., & Soll, J. B. (2007). Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes, 102,* 76–94. http://dx.doi.org/10.1016/j.obhdp.2006.10.002

Lynn, S. K., & Feldman Barrett, L. (2014). "Utilizing" signal detection theory. *Psychological Science*. Advance online publication.

Malle, B. F., & Pearce, G. E. (2001). Attention to behavioral events during interaction: Two actor-observer gaps and three attempts to close them. *Journal of Personality and Social Psychology, 81,* 278–294. http://dx.doi.org/10.1037/0022-3514.81.2.278

Mannes, A. E., Soll, J. B., & Larrick, R. P. (2014). The wisdom of select crowds. *Journal of Personality and Social Psychology, 107,* 276–299. http://dx.doi.org/10.1037/a0036677

MATLAB 8.0 and Statistics Toolbox Release 2012b [software]. (2012). Natick, MA: The Mathworks, Inc.

McNemar, Q. (1969). *Psychological statistics* (4th ed.). New York, NY: Wiley.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review, 115,* 502–517. http://dx.doi.org/10.1037/0033-295X.115.2.502

Moore, D. A., Kurtzberg, T. R., Fox, C. R., & Bazerman, M. H. (1999). Positive illusions and forecasting errors in mutual fund investment decisions. *Organizational Behavior and Human Decision Processes, 79,* 95–114. http://dx.doi.org/10.1006/obhd.1999.2835

Moore, D. A., & Small, D. A. (2007). Error and bias in comparative judgment: On being both better and worse than we think we are. *Journal of Personality and Social Psychology, 92,* 972–989. http://dx.doi.org/10.1037/0022-3514.92.6.972

Otten, S., & Wentura, D. (2001). Self-anchoring and in-group favoritism: An individual profiles analysis. *Journal of Experimental Social Psychology, 37,* 525–532. http://dx.doi.org/10.1006/jesp.2001.1479

Pascal, B. (1962). *Pensées*. London, England: Harvill Press. (Original work published 1669)

Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology, 74,* 1197–1208. http://dx.doi.org/10.1037/0022-3514.74.5.1197

Paulhus, D. L., Harms, P. D., Bruce, M. N., & Lysy, D. C. (2003). The over-claiming technique: Measuring self-enhancement independent of ability. *Journal of Personality and Social Psychology, 84,* 890–904.

Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality, 66,* 1025–1060. http://dx.doi.org/10.1111/1467-6494.00041

Qualtrics Research Suite [Survey software]. (2013). Provo, UT: Qualtrics.

Robins, R. W., & Beer, J. S. (2001). Positive illusions about the self: Short-term benefits and long-term costs. *Journal of Personality and Social Psychology, 80,* 340–352. http://dx.doi.org/10.1037/0022-3514.80.2.340

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15,* 351–357. http://dx.doi.org/10.2307/2087176

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology, 13,* 279–301. http://dx.doi.org/10.1016/0022-1031(77)90049-X

Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science, 3,* 102–116. http://dx.doi.org/10.1111/j.1745-6916.2008.00068.x

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87,* 245–251. http://dx.doi.org/10.1037/0033-2909.87.2.245

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica, 47,* 143–148. http://dx.doi.org/10.1016/0001-6918(81)90005-6

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest, 1,* 1–26. http://dx.doi.org/10.1111/1529-1006.001

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103,* 193–210. http://dx.doi.org/10.1037/0033-2909.103.2.193

Van Veelen, R., Otten, S., Cadinu, M., & Hansen, N. (2015). The cognitive dual pathway model to group identification: A conceptual and methodological integration of self-stereotyping and self-anchoring. *Personality and Social Psychology Review*. Advance online publication. http://dx.doi.org/10.1177/1088868315576642

Wason, P. (1960). On the failure to eliminate hypotheses in a conceptual task. *The Quarterly Journal of Experimental Psychology, 12,* 129–140. http://dx.doi.org/10.1080/17470216008416717

Zuckerman, M., & Knee, C. R. (1996). The relation between overly positive self-evaluation and adjustment: A comment on Colvin, Block, and Funder (1995). *Journal of Personality and Social Psychology, 70,* 1250–1251. http://dx.doi.org/10.1037/0022-3514.70.6.1250

Received January 14, 2015
Revision received July 15, 2015
Accepted July 20, 2015 ■