Self-Enhancement via Redefinition: Defining Social Concepts to Ensure Positive Views of Self

Clayton R. Critcher, Erik G. Helzer, and David Dunning

Cornell University

Running Head: Self-Enhancement via Redefinition

"Words may show a man's wit, but actions his meaning"

--Benjamin Franklin (American author, political theorist, and scientist, 1706-1790)

People live in a world of concrete actions and events, but they think, interpret, communicate, and remember in terms of abstract concepts. People see someone smile, and interpret the action as either *happy* or *ironic*. Someone pats someone else on the back, and they classify it as *friendly, aggressive,* or *ingratiating*. People endorse the idea that they want to achieve *well-being*, but does that state primarily involve healthy relationships with family, more success at work, or just a positive attitude to start the day?

This chapter focuses on the space between the concrete actions that people take and the categories they use to label those actions. It asks how people map the linkages between the concrete experiences they encounter and the abstract concepts they use to categorize, talk about, and remember those experiences. In a sense, we ask how people in their everyday lives complete a task that a psychological researcher often faces. Researchers commonly wish to study interesting concepts such as *intelligence, self-esteem,* or *altruism*, and have to conjure some way to instantiate those concepts in a concrete manner in the laboratory. For example, to instantiate *intelligence*, a researcher might have people complete a math quiz in the lab, tackle a vocabulary test, or examine how many nonsense syllables they remember. More broadly, as Benjamin Franklin indirectly noted in the quote that begins this chapter, researchers must take seriously the mapping between ideas as they exist at a *conceptual level* and how they are to be reflected at an *operational*, or concrete, level.

In this chapter, we make a commonsense observation that people going about their everyday lives also have to make these connections between conceptual and operational levels. But here is the twist. Such mappings are often neither straightforward nor unambiguous. A

single behavior may suggest many different, and reasonable, conceptual mappings. Should a student, for example, who makes a point to talk to the professor after class be considered primarily as *talkative*, *an intellectual*, or just someone infused with *raw ambition*? In addition, a concept might be best exemplified by very different behaviors. When a person is *sophisticated*, does that most directly mean that he or she knows several world languages, has a working knowledge of many different wines, has read a lot of good books, or simply knows how to navigate any social conversation with ease and panache?

To be sure, concepts do contain some core meanings. Ask people how to distinguish the concepts of *intelligence* and *dominance*, for example, and they will show a lot of agreement. They will agree that a large vocabulary has more to do with the former but that asserting one's self in conversation is more relevant to the latter. Classic work in cognitive and social psychology shows that people also agree about many of the core characteristics of what it means to be *in love*, or a *good lawyer*, or even a *chair* (Fehr, 1988, 2004; Rosch, 1975, 1978).

But that classic work also suggests that concepts are fuzzy, that people can disagree about the actions, attributes, and events that best fit the sense of most concepts (for a review, see Horowitz & Turan, 2008), and that the level of disagreement between them can be quite marked. For example, Cantor, Mischel, and Schwartz (1982) asked college students to list the features they associated with such everyday, familiar concepts as *party*, *date*, or being at an *interview*. Ten students provided ten features for each individual concept. On average, 88 features were listed for each concept, but only 15 on average were listed by more than one student. That is, for each concept, 80% of the features each student mentioned was listed by that student alone—indicating more than a fair amount of variability in the mapping each individual brings to mind between abstract concepts and the concrete attributes they link to that concept.

Other work affirms just how variable people's ideas of familiar and important social concepts are. Beck, McCauley, Segal, and Hershey (1988) asked participants to consider such social categories as *quarrelsome* and to rate how much various actions and attributes (e.g., *complaining about having to do someone a favor*) reflected that category. The ratings participants gave were rather stable through time; when participants were asked to give a second round of ratings three weeks after the first, their responses pretty much recapitulated what they had said earlier. Those ratings, however, did not recapitulate what their peers thought. A single person's ratings correlated with another peer's at only the rate of .08 to .50. Fehr (1988) found a similar pattern of disagreement when respondents were asked to consider the concept of *love*. When 96 students were asked to generate attributes of that concept, they generated 183 different attributes, of which only 68 were mentioned by more than one person—and no attribute was mentioned by more than half of respondents. The mean level of agreement between any two of the respondents in this study has been estimated to be only .16 (Horowitz & Turan, 2008). Thus, people agree that love exists, they just do not agree on how it looks, behaves, and feels.

The Aims of This Chapter

This chapter focuses on the fuzzy links between action and idea, and asks what leads people to disagree in their mapping between the concrete and the conceptual. More specifically, we suggest that people tend to shape that mapping in a self-serving way. They tend to endorse concrete-conceptual linkages that place themselves, their characteristics, and their place in the world in a more favorable light. That is, people tend to define such terms as *outgoing*, *self-esteem*, and *tall* in the way that is the most flattering to them.

We review past research showing that people tend to adopt self-serving definitions—i.e., mappings—of social concepts, and that this tendency leads them to reach overly flattering and

much too stable conclusions about themselves that fail to align with objective evidence.. We then consider the recent view, from the psychological and economic literatures, that although people embrace self-serving definitions of social concepts, this does not necessarily lead to self-assessments that are counter-normative. Instead, people can have it both ways, being quite flattering in the way they describe themselves yet still completely accurate in their self-impressions. We then discuss whether the use of self-serving definitions leads to or away from good psychological adjustment. Finally, we explore potential avenues for future research.

The Fuzziness of Social Concepts

A sizable amount of psychology research reveals that the disagreements people display in the linkages they make between the concrete and the conceptual is not random. This work has shown that people appear to take fuzziness in those linkages and exploit them to their own psychic advantage. They emphasize linkages that put themselves in the best possible light—allowing them to claim for themselves positive characteristics while denying negative ones. People mold these linkages in two separate ways. One way we can refer to as the horizontal method. The other is the vertical method.

The Horizontal Method

Social concepts are fuzzy in that is it often unclear which specific attributes or actions best exemplify them. Place an array of concrete attributes in front of people, and they will differ widely in which they pick as most central to any social concept they are quizzed about. For example, given *math skill* and *having a large vocabulary*, people will disagree about which best indicates *intelligence*. Put *diplomatic skill* and *persistence* in front of them, and they will often

disagree about which attribute is more crucial to *leadership*.

But what drives those disagreements? It appears that the motive to think of one's self as a lovable and capable individual fuels some of the difference. For example, when people are asked about the behaviors and characteristics they link to *intelligence*, they tend to emphasize attributes they possess and de-emphasize the ones they do not. People who think they are good at math are more likely to claim that math skill is a central component of intelligence. Those who do not think they have this skill are more likely to claim that other attributes (e.g., having a large vocabulary) are better indicators of intelligence. When it comes to negative concepts, such as *submissiveness*, people reverse. A person who regularly lets their companion choose which movie to go to will probably deny that this behavior is indicative of the negative trait of *submissiveness*; whereas a person who always chooses the movie of the night might very well link a reluctance to do so with the negative trait (Beauregard & Dunning, 2001; Dunning, Leuenberger, & Sherman, 1995; Dunning & McElwee, 1995; Dunning, Perie, & Story, 1991).

These linkages are not merely acts of self-presentation. More automatic measures of the linkages people make show that people connect their personal attributes to good concepts, whereas they avoid linking their shortcomings. Wentura and Greve (2004) asked participants to take a general knowledge quiz with some rather difficult questions (such as *Who wrote* Crime and Punishment*?*). Later, participants were placed in a lexical decision task in which they had to say whether word strings were well-formed words. Of key interest was the speed with which participants could identify the term *erudite* as a word, especially if the word was preceded by a fact that the participants themselves either did or did not know (e.g., Dostoevsky wrote *Crime and Punishment*). When *erudite* was preceded by a fact they knew, they were quicker to label *erudite* as a word than when preceded by a fact they had not known—suggesting that participants

had linked knowledge in that particular area with the abstract concept of *erudite*. This linkage, of course, implies that they themselves are more erudite than not.

The Vertical Method

But even if people agree on the behavioral indicator that best reflects a trait, they still have in hand fuzziness that they can exploit for self-enhancement purposes. For example, let us stipulate, as most people do, that performance on the quantitative SAT test is a central and valid indicator of math skill. There is still some fuzziness left, in that people can differ about how far up the performance ladder on the SAT one must climb to qualify for the trait term *mathematically skilled*. Some people believe that a moderate performance, such as a score of 500 will do; others say that a much more stringent threshold must be met, such as 750 (Dunning & McElwee, 1995).

What drives these disagreements, again, appears to be a desire to maintain positive self-views. People want to claim positive traits for themselves. And once they do, they wish to claim that they *exclusively* have them, relative to others. One sees this activity in how people resolve vertical ambiguity in behavioral indicators. People who score relatively low on the quantitative SAT test, for example, tend to state that a moderate score is sufficient to claim math skill. In doing so, they can claim math skill for themselves. People who score high on the SAT, however, cite a much more stringent threshold. In doing so, they can claim to have math skill and exclude most others in their rather special club, thus heightening the prestige of their own accomplishments (Beauregard & Dunning, 1998; Dunning & Cohen, 1992; Dunning & Hayes, 1996).

The Self as Source of Definitions

Further evidence shows clarifies just how people's definitions of traits and the impressions they have of themselves fall into alignment in a self-serving way. It appears that people often start with themselves and their own attributes and use that information as a guide toward creating a pleasing map between the concrete and the conceptual. That is, people revise their concepts to better match the self they already belief they have. To be sure, at times people may reverse this process and massage their self-concept to better match the versions of concepts they already have in their head) (Kunda & Sanitioso, 1989), but much evidence shows that much of the work people do in maintain self-worth is completed by the redefinition of social concepts .

We know this because altering the self-concept leads to changes in how people define social concepts. One sees this in both horizontal and vertical ways. Concerning the horizontal method, MacDonald, Sulsky, and Brown (2008) asked participants to complete a task that highlighted how much they were either people who were independent versus interdependent on others. Later they were asked to describe their prototype of a good leader. People who had been primed to think of their independent selves tended to describe leadership in a "transactional" way, emphasizing the need for requirements, conditions, rewards, and punishments. Those who, instead, had been reminded of their interdependent characteristics were more likely to emphasize common goals and transcending one's own self-interests. In sum, each group emphasized aspects of leadership that aligned with the particular self that had been primed.

In a similar way, people who succeed at a task respond by raising the vertical standard that must be met to earn a positive social label. People who succeeded at an exercise in which they had to distinguish real from fake suicide notes, relative to those who did poorly, set a higher standard of performance for achieving competence in *social perceptiveness* and *recognizing*

falsehoods (Dunning & Cohen, 1992). In real world settings, college professors who achieve

tenure respond by setting higher standards of productivity and achievement for someone to be

considered worthy of tenure, relative to what they stated before (Eidelman & Biernat, 2007).

The Consequences of Self-Serving Mappings

This self-serving exploitation of fuzziness has consequences, in that it allows people to

hold rosy opinions of themselves that defy the logic of mathematics and reality. However, it also

carries other consequences as well, such as influencing whether people learn from life experience

just how skilled, or not, they are in certain domains.

Unrealistic Self-Views

It is well-known in the psychological literature that people, on average, tend to think of

themselves as anything but average. They tend to think they are more ethical and skilled than

their peers, more likely to attain positive future life events (like long life, for example), while

avoiding negative outcomes (such as being fired from a job). In short, when it comes to positive

attributes, people on average think they are above average; for negative attributes, they believe

they are below average (Alicke, 1985; Alicke & Govorun, 2005 Weinstein, 1980; for reviews,

see Dunning, 2005; Dunning, Heath, & Suls, 2004); but see Chambers & Windschitl, 2004, for a

critique). Such a set of beliefs defies logic, in that it is impossible, except under very severe and

unusual statistical circumstances, for people on average to be "above average."

The self-serving exploitation of fuzziness is a significant source of people's logic-defying

self-descriptions in that they report such descriptions mostly when dealing with characteristics in

which there is ambiguity to exploit but not when such ambiguity is unavailable. For example,

traits such as *sophisticated* and *talented* are quite broad and allow for much cherry-picking in the concrete actions and attributes that most reflect them. As a consequence, people are more likely to claim they are above average for these traits. For other traits that are constrained in meaning—that is, that are less ambiguously definable (e.g., *neat, thrifty*)—people show little, if any, bias (Dunning, Meyerowitz, & Holzberg, 1989; Suls, Lemos, & Stewart, 2002).

Other researchers similarly show that the exploitation of fuzziness leads to self-serving beliefs. Hsee (1996) asked participants to take a 20-item quiz, and at the end told them that all the correct answers were "A". For some participants, he then said that he would pay them for each odd-numbered item they got right. Participants in this case provided rather unbiased reports of their performance on the quiz. In the other condition, the numbers next to the questions were replaced by figures that looked, more or less, like the Chinese characters for *yin* and *yang*. These participants were told that they would be paid for their performance on the 10 items attached to the *yin*-like figures, but what looks yin-like, however, is somewhat fuzzy. Thus, it is not a surprise that participants over-reported how many items they got right, defining yin- and yang-ness in such a way to bolster their prospects. The same happens in the real world in regards to self-descriptions of physical attractiveness. Relative to how others rate them, men and women rate themselves higher on fuzzy concepts such as *attractiveness* and *sexiness* than on a more well-defined attributes like *body size* (Donaghue & Smith, 2008).

Stability of Self-Concepts

Self-serving definitions may be at the heart of another finding in psychological research: that people maintain fairly stable views of themselves despite the fact that they receive feedback that is quite variable and inconsistent with those views (e.g., Mazar, Amir, & Ariely, 2008). For

example, medical students' self-rated ability at the end of medical school is strongly correlated

with their self-ratings during their first year, despite the fact that their self-ratings at the end fail

to correlate with either supervisor ratings or objective examination scores (Arnold, Willoughby,

& Caulkins, 1985). Moreover, the fact that these self-ratings correlate so strongly ($r = .54$)

seems to imply insensitivity to the vast amount of new information that is surely garnered after

hundreds of hours of clinical training. On our account, these self-views are maintained through

redefinition. Medical students in this study likely redefined their criteria for *medical ability* in

order to accommodate their performance feedback.

In general, when people engage in behavior that is relevant to their standing on a trait,

they stand to receive three classes of feedback: they may learn that their believed standing on a

trait is calibrated and accurate, they may learn that their believed standing on a trait is overblown

and unrealistically optimistic, or they may learn that their believed standing on a trait is better

than they originally thought. All types of feedback may lead to self-serving redefinition. For

example, when people receive information from the world that indicates that they have exceeded

even their own idiosyncratic standards of a desirable trait, there is evidence that they will revise

their trait definition to reflect this improvement in the self. Upon receiving an A, Tom may

update his definition of a "smart" student from one who maintains a B+ to one who holds at least

an A- average (see Eidelman & Biernat, 2007).

Conversely, when people receive unflattering feedback from the world, they can

barricade their positive self-views by tagging the new information as non-diagnostic of the

threatened trait. Greve and Wentura (2003) have showed this most directly. In one of their

studies, participants competed against a confederate in four trivia quizzes. The competition was

rigged such that the participant outperformed the confederate in two domains (e.g., politics and

fine arts), but performed relatively poorly in two domains (e.g., history and natural sciences). When asked in what domains a person with good general knowledge would excel, participants saw the domains at which they had been randomly assigned to excel as more diagnostic of a person's general knowledge.

This process, known as *self-immunization* is not merely a product of conscious, effortful rationalization. In Wentura and Greve (2004), implicit measures showed that people possessed automatic connections between specific pieces of concrete information and flattering abstract traits those concrete tidbits might reflect, but only to the extent that the potentially flattering concrete information was true of the self. This process, it should be noted, functions on the basis of two principles. First, people must be aware of their strengths and shortcomings. As such, this self-enhancement technique stands in contrast to lower level distortions that might block people's awareness of their failures. Second, people must adjust their criteria for a desirable trait so as to exclude the domain in which they have experienced failure.

Self-enhanced trait definitions are also supported by reliable differences in the way people conceptualize successes and failures in everyday life. Kurman (2003; Study 1), for example, showed that people contain failures in part by dissociating them from their more global sense of self-worth. When participants judged a list of 30 hypothetical successes and 30 hypothetical failures in terms how global versus specific each is, ratings for successes reflected more global construals than did ratings for failures. In addition, participants rated how much of a self-esteem boost or hit they would receive for each of the 60 events, were they to happen. The global versus specific construal of the event was found to moderate the relationship between failure and self-esteem: the more specifically participants construed failures, the less they anticipated a hit to their self-esteem. In addition, work by Taris (1999) suggests that not only do

people represent their successes more abstractly than their failures, but when spontaneously

describing their actions, they cite more instances of success than failure.

In sum, people may only infrequently face the burden of reconsidering their standing on

flattering traits when confronted with feedback. People describe failures less frequently and rope

them off as specific instances that only loosely implicate more global self-worth, allowing them

to hold them independent of their broader, positive traits. When feedback (either negative or

positive) suggests a mismatch between people's idiosyncratic trait definitions and trait-relevant

behavior, they can make affordances for behavioral data by adjusting their criteria for desired

traits. If feedback is positive, people can adopt more selective criteria for a desired trait; if

feedback is negative, people may tailor their definitions so as to prune away domains in which

failure has occurred.

Are Self-serving Social Concepts Normative?

At first blush, the use of self-serving definitions would appear to be counter-normative, in

that they lead to judgments that are mathematically impossible and cause them to ignore

informative feedback that the world is supplying to them. Traditionally, psychologists have

viewed self-enhancing self-views as directly antagonistic to accuracy. Taylor and Brown (1988),

for example, contrasted the self-enhancement of the mentally healthy with the realism of the

depressed. Noting that many people who claimed to be above-average may not have been

exaggerating their self-views, other researchers did more to differentiate "true self-enhancers"

from those who were merely accurate in their above-average assessments (John & Robins, 1994;

Kurt & Paulhus, 2008; Kwan, John, Kenny, Bond, & Robins, 2004; Paulhus, 1998). From this

perspective, self-serving social concepts are normatively unjustifiable to the extent that they

offer flattering self-views that depart from reality.

We feel compelled, however, to point out that this analysis may be inaccurate and premature. There is a way in which people may self-enhance, even severely, yet still have judgments that are accurate and normative. Because self-serving trait definitions essentially connect two levels of information—concrete details about one's life and abstract inferences about one's standing—it may be possible for people to hold overly rosy views of themselves at one level while maintaining perfect accuracy at the other.

Specifically, according to Schneider (2001), there is a telling difference between *knowledge*, concrete facts about ourselves and our environments, and *meaning*, the abstract significance granted to such details. The nature of concrete knowledge is more exact, simpler to describe, and offers a clearer normative standard against which to compare one's self-knowledge. A person either received an A on her economics exam or did not; can dunk a basketball or not; has a clean driving record or not. In each case, concrete reality is precisely defined, and it is relatively simple to assess whether one's understanding of the world is accurate or deluded.

Schneider (2001) pointed out, as do we, that one has great latitude in attaching abstract meaning to one's concrete knowledge. Although the B atop one's economics exam or the silver medal around one's neck leave little room for interpretation, rational people may differ in concluding whether or not the mark reflects an aptitude for economics or whether the quality of judging in the figure skating competition was poor. In this way, people may be perfectly accurate their concrete knowledge ("I received a B on the exam") but put a self-flattering spin on this knowledge in the realm of meaning ("Given I'm just a freshman, I'm practically an economics genius!"). Economists, for example, have recently asserted this analysis of the above

average effect, and have proposed that the flattering judgments that people make of themselves are perfectly appropriate—and not at all remarkable—if based on self-serving definitions of social traits (Santos-Pinto & Sobel, 2005).

Of course, there is some point at which self-serving definitions move from exploiting ambiguity to defying reason (e.g., a repeat-offending felon seeing himself as "law-abiding"). Although acknowledging that there are some constraints on how much one may reasonably twist a social concept's meaning, our point is that for many social concepts there is great latitude in how reasonable people may (and do) differ in their concrete definitions of those concepts. Given this, there is plenty of room for people to exploit these honest disagreements to their advantage, without necessarily being biased.

Greve and Wentura (2003), in their work on *self-immunization*, have offered a clear demonstration of how people can accurately accept their limitations at the concrete level by giving those limits a flattering spin at the conceptual level. Recall that participants took quizzes, roughly half of which they excelled at and the other half they did not. Afterwards, the participants offered accurate judgments of their own skill in the tested domains. Their sense of relative competence in each tested domain closely aligned with their performance relative to the confederate. But this accurate, concrete knowledge was complemented by a flattering assessment of what this knowledge meant: Participants emphasized the areas they had excelled at as being the more crucial components of "general knowledge."

Dunning and Beauregard (2000) have shown similar reactions when people are given feedback about themselves. In their study, they gave participants false personality feedback about their social competence. Those who started off with a low opinion of themselves incorporated this feedback into their self-views. In contrast, those who were already confident of

their own social skills dismissed the relevance of the feedback, continuing to believe they were socially competent even when one concrete datum told them otherwise. To the extent that these initial self-views were accurate this form of self-immunization is not only justifiable, it is probably warranted. There is a massive caveat to this observation, however. The accuracy of self-views is notoriously poor, see Dunning, 2005; Dunning, Heath, & Suls, 2004), suggesting that people, on balance, may do better to pay attention to the feedback they receive from the world, no matter how unpleasant, rather than to dismiss it as uninformative.

But a Step Too Far: Confusing Knowledge and Meaning

Thus, there is a way in which people can have their evaluative cake and eat it, too. They can possess both accurate knowledge and a pleasing image of the self. However, achieving this state depends on a crucial assumption--that people can keep straight what is meaning and what is knowledge, so as not to confuse the abstract opinions they hold about themselves with the facts at the concrete level. More specifically, enhanced views of the self can only be normative if people remain aware that self-flattering definitions of social concepts are merely their subjective views and not facts in themselves. Other people need not agree with them, nor should their personal definitions in any way be construed as universal. For example, if Paul has a rather Paul-like definition of intelligence, but understands that Art could just as reasonably construct his own different definition of intelligence—one based on Art's idiosyncratic skills and shortcomings—then Paul's self-serving definitions need not lead to counter-normative results.

However, when people take the additional step of confusing meaning and knowledge, beginning to believe that their abstract self-beliefs are not just opinion but rather the way things really are (e.g., *I am* objectively *more intelligent*), people and their decisions may begin to depart

from a purely normative position. Such a departure can occur in two inter-related ways. First, people may decide—as fact—that they really are better than other people when it comes to social traits. Second, they may decide that the definition they have created represents the only right and true one. In a sense, we argue that self-serving definitions become problematic when people commit the error of *naïve realism*. Naïve realism is the mistake of tacitly assuming that one's opinions and beliefs represent the true state of the world and the way it should be. It is to mistake subjective opinion with objective truth (Ross & Ward, 1996).

Evidence suggests that people pervasively take this step, considering their definitions to be *the* definitions that should be used in evaluating social life, thus leaving the realm of normative judgment. They apply their definitions universally, using roughly the same definitions in judging other people as well as themselves (Paulhus & Landolt, 2000). For example, McElwee, Dunning, Tan, and Hollman (2001) asked participants to rate the leadership ability, creative talent, and intelligence of a number of target individuals. They found that participants judged others according to the same (flattering) criteria that they used to judge themselves (see also Dunning & Cohen, 1992; Dunning & McElwee, 1995; Dunning, Perie, & Story, 1991). When a target was a task-oriented individual, participants rated that target's leadership ability more favorably if they themselves were task-oriented individuals. Furthermore, the influence of participants' self-views on judgments of the targets was fully mediated through participants' self-serving trait definitions as explicitly expressed. This suggests that participants went beyond putting a flattering spin on the truth ("That I know how to write music means I am a creative person") to insisting that their spin actually was the truth ("People who can't write music aren't creative people."). Put another way, problems arise when people move beyond saying that concrete abilities they possess are *sufficient* for concluding that they possess valued abstract

qualities, to concluding that it is *necessary* that someone possess these concrete abilities if they wish to be accorded the positive abstraction.

Further evidence that people confuse their subjective opinions with social reality comes from data on interpersonal disagreement. For example, therapists agree only modestly ($r = .23$) with their clients about their own therapeutic effectiveness (Bogwald, 2001), and managers and employees show very little agreement about the former group's influential ability (Blickle, 2003). To be sure, these disagreements are due to a number of psychological and social factors; however, when people differ in their evaluation of one another, a good share of that disagreement can be traced to the parties' differential and perhaps self-serving resolution of trait ambiguity (Hayes & Dunning, 1997). Story (2003) found that the greatest disagreement between self and other on trait ratings was among participants who differed in their idiosyncratic definitions of a given trait, and that inter-rater agreement could be improved by providing peer raters with the target's own definition of a given trait. Notwithstanding, ambiguity alone may not be sufficient for disagreement. After all, independent 3rd party raters can show remarkable agreement about a target's standing on a number of ambiguous traits, even on the basis of limited information.

Equally problematic is that the spin that people put on concrete knowledge can interfere with their memory for that knowledge. For example, participants who have just been led to believe that either extroversion or introversion was a trait that reflected an aptitude for future success became more likely to recall past behaviors that were consistent with the desired trait (Brunot & Sanitioso, 2004), reported it was metacognitively easier to recall such behaviors (Sanitioso & Niedenthal, 2006), recalled just-received feedback to be more consistent with the desirable trait (Sanitioso & Wlodarski, 2004), and judged recent behavior to be more in line with

the desired characteristic (Sanitioso, 2004). Those who were told that frequent tooth brushing was actually harmful to dental hygiene recalled brushing their teeth less often (Klein & Kunda, 1993). To the extent that the positive spin people place on self-relevant knowledge then distorts people's recall of concrete details about the self, self-serving conceptual definitions have set in motion a self-perpetuating process that pushes one's self-views from merely sympathetic to indisputably incorrect.

Finally, when flattering spin is confused with positive knowledge, people will be led astray if they use this flattering spin in trying to make predictions about their concrete future. For example, people seem to use self-serving criteria in evaluating their morality, giving weight to praiseworthy intentions, even when those don't actually translate into praiseworthy behavior (Kruger & Gilovich, 2004). When people rely on that positive spin in judging the likelihood that they will engage in future moral behaviors, they offer predictions that are overly optimistic. Across many demonstrations, Epley and Dunning (2000, 2006) found that people were much too optimistic in predicting their own future prosocial behaviors, such as voting and giving to charity (see also Balcetis & Dunning, 2008), whereas they were much more accurate in predicting the behavior of their peers.

Summary

Much of people's representation of the world is inherently subjective, and people's beliefs are not outside the bounds of rationality when they choose to weave a hopeful interpretation of the facts before them. This tendency can have the pragmatic benefit of allowing people to accept unflattering concrete details about themselves while putting an esteem-buffering spin on such shortcomings. Furthermore, the tendency can lead people to recognize aberrant

negative feedback for what it is—aberrant.

But people run into trouble when they fail to recognize the subjectivity of their own perspectives on the world. Gestalt psychologists demonstrated that people effortlessly move beyond the concrete, perceptual details of a stimulus and automatically extract out a reality that, in truth, exists only in the mind of the perceiver (Köhler, 1930). The visual system is not "wrong" in offering this perspective on reality, but it is ignorant in implicitly assuming that what it sees *is* reality. As social perceivers, people seem to operate under a similar type of naïve realism—believing that their perspective on social reality *is* reality (Icheiser, 1943; Ross & Ward, 1996). In this section, we have seen that people are naïve realists in their approach to social concepts. They do not see their own self-enhancing perspective on the world as one theory among many, but instead see it as *the* theory of the social world.

Are Self-Serving Definitions Adaptive?

However, even if people use self-serving definitions in a non-normative way that leads to inaccuracy, they might still be using them in a way that is adaptive—that is, in a way that leads them to beneficial rather than costly outcomes. The psychological literature has long harbored a controversy about whether self-flattering self-views are adaptive or maladaptive. Taking issue with the consensus among clinicians, Taylor and Brown (1988, 1994) argued that positive illusions do not interfere with, but instead provide, the foundation for psychological health. However, subsequent research cast doubt upon these conclusions. Instead of being unconditional models of psychological adjustment, self-enhancers have been observed to be heavily narcissistic (Gosling, John, Craik, & Robins, 1998; Robins & Beer, 2001), defensive (Colvin, Block, & Funder, 1995), and less well-adjusted in the eyes of mental health professionals (Robins & John,

1997; but see Taylor, Lerner, Sherman, Sage, & McDowell, 2003).

We think that to address this controversy completely, one must consider not only the fact of self-enhancement but also the avenue by which people achieve it. The fact that self-serving definitions underlie so many positive illusions suggests that a complete treatment of the adaptiveness issue must specifically address the types of illusion prompted by self-serving definition. We first introduce two relevant distinctions and discuss them in turn—the process responsible for bringing about self-enhancement and the nature of the adaptiveness measure. We conclude by alluding to two additional distinctions that are important to the adaptiveness question.

The Cause of Self-Enhancement.

Self-enhancement may occur for a variety of reasons. People may harbor enhanced self-views because they are just not skilled or knowledgeable enough to know how incompetent they are (Dunning, Johnson, Ehrlinger, & Kruger, 2003; Ehrlinger, Johnson, Banner, Dunning, & Kruger, 2008; Kruger & Dunning, 1999); because the ambiguity of certain concepts allows people to draw self-serving conclusions that seem reasonable (Dunning, Meyerowitz, & Holzberg, 1989; Felson, 1981); because negative feedback is often couched in ambiguous or euphemistic terms (Goffman, 1955); because other people feel social pressure to provide feedback that validates a person's positive self-conceptions (Swann, 1983; Taylor & Brown, 1988); or because one wishes to motivate oneself to persevere in implementing an action plan (Taylor & Gollwitzer, 1995).

One can ask if the adaptiveness of self-enhancement differs depending on the specific psychological pathway that produced it. On the one hand, it might seem strange to think that the

adaptiveness of self-enhancement would depend on the process and circumstance that produces the illusion. After all, the likelihood of contracting a disease from unprotected sex does not change based upon whether one engages in it due to ignorance of condoms' protective function or motivated reasoning that such behavior is safe. On the other hand, to the extent that different pathways to self-enhancement covary with other features of the situation and the person, the adaptiveness of self-enhancement might depend upon how the self-enhancement arose. Some pathways to self-enhancement may arise via processes in which they confer functional advantages. For example, deliberately choosing to be optimistic while implementing an action plan may help one to persevere when "the going gets tough." Here, a self-enhancing bias would be a conscious strategy, and one knowingly chosen to produce adaptive results. But, on the flip side, fate may not be so kind if self-enhancement follows from a simple but honest failure to recognize the limits on one's own ability, due to inexperience. The seafloor is littered with airplane fragments flown by too many inexperienced, and thus inadvertently overconfident, pilots.

In addition, the different reasons for the emergence of self-enhancement have different implications for whether distorted self-knowledge is correctible. For some causes of self-enhancement, people may be able to shift to more accurate self-understanding if the situation calls for it. By eliminating the definitional ambiguity of a quality on which one will judge oneself, or by reverting from a mindset focused on implementing an action plan to a mindset that deliberates about a proper course of action, one may be able to impose "reality checks" that allow for accurate perceptions when they prove crucial. In contrast, poor metaknowledge about how to reasonably evaluate the self, or the frequent receipt of distorted feedback, may produce misperceptions that are uncorrectable (see Taylor et al., 2003, for a similar argument).

With this in mind, we have to admit it is likely that there is not a straightforward answer to whether self-enhancement through conceptual definition is adaptive or not. At times, people may benefit from the positive illusions that these definitions support; at other times, these illusions may prevent people from spotting errors when they occur. And even though bias can be corrected to some extent by providing people with specific definitions upon which to base judgments of themselves (Dunning, Meyerowitz, & Holzberg, 1989), it is unclear whether such techniques actually lead people to more realistic self-assessments or merely reflect temporary accuracy that stems from the compliance of experimental participants.

The Nature of Adaptiveness.

In judging whether self-enhancement is adaptive, one must be careful to consider the different faces of adaptiveness. For example, it may be useful to distinguish *internal* adaptiveness from *external* adaptiveness. Internal adaptiveness refers to intrapsychic benefits such as emotional resilience, a positive sense of self-worth, and a hopeful outlook on the future. External adaptiveness refers to benefits in the person's attempts to interact with the world, such as interpersonal harmony, positive or worthy behaviors, and goal achievement. In reviewing the literature, one finds clearer evidence that self-enhancement promotes internal adaptiveness than external adaptiveness.

Internal Adaptiveness. Self-enhancement encourages a positive sense of self. Kwan, John, Kenny, Bond, and Robins (2004) found that "self-specific" self-enhancers—those who had a distorted positive view of the self, but not simply a distorted positive view of everyone—had higher self-esteem. Kurt and Paulhus (2008) found that such individuals saw themselves as higher in personal adjustment: They worried less in life and felt like life had not offered them a

raw deal. But curiously, their peers seemed to have the opposite opinion. Finally, self-enhancers maintain greater emotional equanimity in the face of tragedy (Bonanno, 2004)—rated by mental health professionals as better adjusted (Bonanno, Field, Kaltamn, 2002) and possessing fewer biological markers of stress (Bonanno, Rennicke, & Dekel, 205).

Self-enhancement through conceptual re-definition may confer internal adaptiveness in a qualitatively different way as well. Roese and Olson (2007) have argued that people stand ever vigilant against threats. When people detect a threat, a surge of negative affect alerts them that there is a problem that must be addressed. The aversiveness of the affective experience motivates a quick threat response. If people do not respond to the threat relatively quickly, then their affect-based threat detection system will not return to baseline, and they will not be ready to notice new threats as they appear. From this perspective, even if self-enhancement through definition is at times maladaptive from a local perspective, such a strategy may promote adaptive vigilance when considered from a broader perspective.

External Adaptiveness. What about external adaptiveness? Although there is some evidence that self-enhancers have more positive interpersonal relationships, this benefit seems to be limited to those self-enhancers who have unjustifiably positive views not only of themselves but of other people as well (Kwan et al., 2004). Outside of this subsample, though, self-enhancers often elicit negative reactions from others. Others may see self-enhancers as condescending (Colvin, Block, & Funder, 1995), disdainful (cf. Gibson & Oberlander, 2008), and filled with too much braggadocio (Paulhus, 1998).

Looking more narrowly at self-enhancement through definition, a great deal of disagreement can arise when people bring their own self-enhancing definitions to an interpersonal context (McCrae, Stone, Fagan, & Costa, 1998; Story, 2003). One area where

disagreements may be especially likely to produce rancor is in the domain of morality, where ambiguity about what is a correct course of action is at its highest. People differ not only in their moral convictions, but in their beliefs about whether an issue even falls within the moral domain (Bauman & Skitka, 2009). For example, although liberals believe that moral questions deal with harm or fairness, conservatives have a more expansive definition that includes in-group loyalty, purity, and respect for authority (Graham, Haidt, & Nosek, 2009). Such divergences shed light onto why conservatives and liberals often seem not only to differ in opinion, but also have fundamentally incompatible perspectives on contentious issues like same-sex marriage. Although these definitional differences may stem only in part from self-enhancement (e.g., a loyal ethnic-group member may want to believe his allegiance is a moral issue), this demonstrates nonetheless the way divergent conceptual definitions can produce discord.

Finally, self-enhancement can lead one to engage in maladaptive behaviors. In a series of studies, Camerer and Lovallo (1999) showed how people's inflated self-views lead them to make unwise economic decisions. For example, when participants were told that their economic outcome in a game would be determined by their own skill (as opposed to luck), participants were more likely to unwisely enter the game and lose money. Fortunately, there are often social checks on one's own (idiosyncratically defined) delusions of grandeur. Even if one has defined one's own skills in an overly flattering way, outside observers (e.g., a boss) can keep self-enhancers in check. Dunning, Heath, and Suls (2004) noted that, problematically, those who have the power to make the most impactful decisions are often those who are so powerful that they have few social constraints on their decision making. In a business, this person would be the CEO. Reflecting CEO overconfidence, decisions to acquire other companies rarely meet with enthusiasm in the stock market (Malmendier & Tate, 2003). And as those playing the

market predicted, the most overconfident CEOs' acquisitions ended up offering the most

negative returns on their investments (Hayward & Hambrick, 1997).

Summary

In summary, the question of the adapativeness of self-enhancement through conceptual

definition is not a clear-cut issue. In this section, we have considered two distinctions that may

be relevant in answering this question: the cause of self-enhancement and the nature of the

adapativeness measure. We suspect that still more distinctions may be important in acquiring a

nuanced understanding of the adapativeness of self-enhancement. For example, exploiting the

ambiguity in conceptual definitions may offer short-term benefits at the expense of long term

costs (Bonanno et al., 2005; Paulhus, 1998; Robins & Beer, 2001). In addition, even when

exploiting conceptual ambiguity to bolster one's self-image does not have maladaptive

consequences for self-enhancers, it may adversely impact those around them (e.g., Effron,

Cameron, & Monin, 2009). Although a variety of research has studied in isolation the

distinctions introduced in this section, future research should examine how these orthogonal

distinctions interact in conferring adaptive benefits.

New Questions

Faced with this extensive body of literature, one may wonder whether there is anything

new to say, empirically at least, with regard to self-enhancement through definition. We believe

there is, and focus on two potential lines of investigation that could expand an understanding of

the role of conceptual re-definition in social life. The first issue is the scope of social concepts

that people might paint in rosy, self-serving tones. A second might explore the way in which

collectives, ranging from dyads to entire cultures, agree upon definitions that flatter the group

rather than the individual.

The Scope of Self-Serving Definitions

The work described above, although far-ranging, has tended to focus on only one

category of social concepts that people might re-define in their favor. That category consists of

social traits, such as *intelligence* or l*eadership*. There are a variety of other social concepts,

equally full of ambiguity, that people might link to concrete behaviors in an attempt to make

sense of themselves in the social world. Let us enumerate just a few.

Goals. "I want to get healthier," cries the downtrodden Northeasterner, emerging from a

long winter of pizza, beer, and pork rinds. What does he mean by that? The goal itself is an

ambiguous one, and we suggest that to resolve this ambiguity our winter sloth (and people in

general) will call upon self-serving construals of "getting healthy" so as to evaluate progress

toward the goal in the most flattering way. Although this idea has intuitive appeal, we know of

no direct research that examines this possibility.

There are at least three ways to capitalize on the ambiguity of goals so as to reach a self-

serving conclusion. First, a person can reevaluate the *how* of the goal— how (through what

concrete behaviors) does one make goal progress? Second, a person can re-construe the *what* of

the goal— what is the target that indicates goal completion? Third, a person can exploit the

when of the goal— when should progress be measured: at short-term behavioral

accomplishments or long-term end-state completion?

When self-enhancing through the *how* of their goals, people can take a horizontal

approach, by including and excluding types of behaviors from the definition of the goal, or a

vertical approach, by redefining the amount of a behavior that is tied to a goal. Whether taking the stairs to a third-floor office constitutes "getting healthy" probably depends upon whether or not the walker made good on his promise to go to the gym that day. If he didn't, he may temporarily include this somewhat trivial exertion as an indicator of his goal progress, even though he might overlook the very same behavior if he had more convincing concrete evidence to link to his abstract goal of "getting healthy." Similarly, if "getting healthy" is inextricably linked to the concrete behavior of going to the gym, one may still capitalize on vertical ambiguity by redefining the *amount* of gym attendance necessary: is getting healthy a matter of going to the gym three times a week or five times? If a person only goes once, but works up a "really good sweat," that behavior, too, might a serve as a sufficient self-serving indicator of progress toward the goal of "getting healthy."

Alternatively, people may render self-enhanced visions of goal progress by keeping the behavioral indicators constant while altering the *what* and *when* of goal completion. Achieving the goal of "getting healthy" could mean losing 15 pounds or 50; improving muscle strength or cardiovascular health; eating better or being more active. Thus, people may believe they have reached the goal of getting healthy by citing one of a number of different end-states that are potential *what*s of goal fulfillment. In much the same manner, people may capitalize on the ambiguity of the *when*s of goal fulfillment: Is each trip to the gym its own success or does one only achieve the goal of getting healthy after establishing a long-standing work-out routine? The answer to this may depend on what concrete changes the goal-setter was actually able to accomplish.

Attitude Objects. In the quest for positive self-regard, people may find it desirable to claim attitudes that signify them as a person of worth (see Katz, 1960). It is well known that

people feel better about themselves when they adhere to culturally-accepted beliefs and practices

(e.g., Solomon, Greenberg, & Pyszczynski, 1991); so, for example, a college freshman might

wish to hold attitudes that are left of political center in order to garner favor with his professors,

and a guest at a party may wish to claim attitudes that are sophisticated in order to be accepted

by the *nouveau riche* in attendance. However, in order to render this attitude construction

something more than complete self-deception, people must link abstract attitudes to actual

attitude objects.

According to attitude representation theory (e.g., Lord & Lepper, 1999), people generate

attitudes on-line by calling to mind particular exemplars of the attitude category. When a person

offers her attitude about politicians, for example, she uses her evaluation of an accessible

relevant exemplar (*Barack Obama*) as a cue toward her feelings toward politicians as a general

category ("politicians are agents of change"). Put another way, people define their attitudes on

the basis of concrete examples that are stored away in memory. Of course, for most attitude

domains, people have a wealth of material from which to draw, and this fact suggests that the

product of such a search can be a self-enhancing attitude. This road toward self-enhancement

can happen entirely automatically, and without awareness. When the undergraduate student

from above is asked for his opinion on affirmative action, the political science classroom in

which the question is posed, or the student sitting next to him at the moment of questioning, may

prime certain exemplars in memory, and he may, as a result, produce exactly the attitude that

bolsters his sense of self-worth as a broad-minded and enlightened college student (*cf.* Lord,

Desforges, Fein, Pugh, & Lepper, 1994).

Intentionality. When people form links between a concrete behavior and an abstract trait,

their inferences are often driven by their analysis of the actor's motivations. Whether "kicking a

dog" is represented as "mean-spirited," "clumsy," or "unfortunate" depends upon whether the actor performed the behavior with intent. But intentionality, like other social concepts, is an ambiguous feature of behavior. Because people have very little access to the underlying motives for others' (or even their own) behavior, self-serving motivation can, and does, run rampant upon people's attributions of intentionality, promoting self-flattering links between behaviors and traits.

Past work in moral psychology, for example, has shown that people can arrive at desired conclusions about an actor's moral character by deeming that act *intentional*. As illustration of this, consider a scenario used by Knobe and colleagues (Leslie, Knobe, & Cohen, 2006), in which participants read about a CEO who learns that a new corporate policy will increase profits, but will carry the side-effect of harming the environment. The CEO responds to the potential negative consequence by saying, "I don't care at all about harming the environment. I just want to make as much profit as I can." Note that this concrete statement could be linked to the trait of "shrewd businessman" or "heartless capitalist;" and could be interpreted as "looking out for the interests of the company employees" or "being greedy at any cost."

Nonetheless, when drawing a link between concrete behavior and abstract meaning, people are constrained by whether the behavior is an intentional one—after all, it doesn't make sense to say that the CEO's acts are *greedy* if he did not intend to do them. As a result, when participants are asked to evaluate the CEO's behavior, not only do they offer unwavering condemnation of his moral character they believe that he harmed the environment intentionally.

One route by which people can see intentionality where it may or may not exist is by shifting their definition of *intentional* so as to include the actor's motivation at hand. In a replication of the above study, Tannenbaum, Ditto, and Pizarro (2008) compared the reactions of

environmental absolutists, who are unwilling to make pragmatic environmental trade-offs, and environmental relativists, who will consider pragmatic trade-offs for a greater good. As expected, rabid environmentalists offered harsher condemnation of the CEO than did the more moderate environmentalists. When asked to support their judgments, those who condemned the CEO most cited a "strict" definition of intentionality as their standard of blame, one that included any sort of foreknowledge that bad consequences might be incurred; whereas those who tempered their moral outrage cited a "loose" definition of intentionality, claiming that the CEO flatly stated that he held no preexisting goal to hurt or harm the environment (Tannebaum et al., 2008; cited in Ditto et al., 2009).

In sum, then, if preserving warm thoughts of the self requires that a person render moral condemnation upon another actor, people can lower their standard for an intentional act to one that encompasses more mundane motivational states, such as negligence or recklessness. One might also call upon this lowered standard when accepting praise for one's own ambiguously intentional behavior. Conversely, if self-enhancement goals are served by removing blame from one's self, one can shift to a stricter standard of intentionality, one that rules out these very same motivational states.

Social Definition at the Collective Level

In recent years, self-enhancement researchers have become increasingly interested in whether members of individualist and collectivist cultures differ in their patterns of self-enhancement (Balcetis, Dunning, & Miller, 2008; Heine, Lehman, Markus, & Kitayama, 1999; Sedikides, Gaertner, & Toguchi, 2003). To date, the focus of the debate has been on whether people from collectivist cultures enhance their own standing on positive traits to the same degree

that individualists do. A different perspective on enhancement, though, begins with the more

general observation that groups of people, from dyads to entire cultures, maintain positive

feelings about their collective, and motivate behavior within the collective, by defining abstract

positive traits in ways that map onto group behavior. Looking at the ways in which groups

maintain their status by engaging in enhanced trait-definition is an avenue for a great deal of

fruitful research.

One way to maintain unwavering positive feelings toward shared social groups is through

group-level definitions of valued concepts. Work by Aloni and Bernieri (2001), for example,

suggests that romantic couples define "love" according to the behavioral patterns that are most

prevalent in their own relationship. Just as self-serving definitions of one's own traits can be a

source of great interpersonal disagreement, these relationship-based trait-definitions blind

couples' social judgments: partnered people are less accurate than unpartnered people at

detecting romance in couples whose behavioral patterns differ from their own (Aloni & Bernieri,

2001). In fact, one may look at the largest collectives (religious sects, political parties, or

nationalities) as being fundamentally bound by collective agreement upon definitions of

ideological concepts. Abstract notions of desirable concepts such as *freedom*, *faith*, or *fidelity*

are likely to be represented as patterns of behavior that are common to the group that holds them.

As an example, the constellation of behaviors and mental states that constitute *happiness*

appears to be highly dependent upon the culture from which one emerges. In a series of studies,

Tsai and colleagues (Tsai, Knutson, & Fung, 2006) showed that children and adults from Asia

link the abstract emotion of happiness with the concrete experience of low-arousal positive

affect, while European children and adults link the same emotion with high-arousal positive

affect. An analysis of children's books and sacred texts from both cultures further illustrated this

culturally-specific linking: European culture portrayed the trait of happiness as a high-arousal experience, whereas Asian culture portrayed the same trait as a low-arousal experience (Tsai, Louie, Chen, & Uchida, 2007; Tsai, Miao, & Seppala, 2007). In light of other research suggesting that from birth Asian babies show less arousal than do their European counterparts (Freedman, 1974; for a review, see Rushton, 1999), one could reasonably explore the possibility that shared collective definitions of *happiness* and other virtues are biased so as to favor the collective's typical patterns of concrete behavior.

Concluding Remarks

People desire to think positively of themselves, and fortunately for them, the inherent ambiguity of the world often allows people to define social concepts in self-flattering ways. In this chapter, we have highlighted the creative strategies people employ in order to "get the most" out of the abstract meaning they assign to concrete behavior. However, people's desire to paint themselves in a positive light is sometimes at odds with other important needs—the need for positive relationships with other people, the need to foresee events of personal consequence, and the need to act in ways that keep the self safe and secure. When people's self-enhancing definitions of the world become too divorced from, or become confused with, concrete reality, they may have difficulty satisfying all of the above needs.

Thus, it may be true that the world looks more beautiful through rose-tinted lenses. And although this tinted world is not imaginary, the challenge may be for people to recognize, at some level, that it is not really real.

Acknowledgements

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.

Aloni, M., & Bernieri, F. (2004). Is love blind? The effects of experience and infatuation on the perception of love. *Journal of Nonverbal Behavior*, *28*, 287-295.

Arnold, L., Willoughby, T. L., & Calkins, E. V. (1985). Self-evaluation in undergraduate medical education: A longitudinal perspective. *Journal of Medical Education*, *60*, 21-28.

Balcetis, E., & Dunning, D. (2008). A mile in moccasins: How situational experience reduces dispositionism in social judgment. *Personality and Social Psychology Bulletin*, *34*, 102-114.

Balcetis, E., Dunning, D., & Miller, R. (2008). Do collectivists know themselves better than individualists? Cross-cultural studies of the holier than thou phenomenon. *Journal of Personality and Social Psychology*, *95*, 1252-1267.

Bauman, C. W., & Skitka, L. J. (2009). In the mind of the perceiver: Psychological implications of moral conviction. In D. Bartels, C. W. Bauman, L. J. Skitka, & D. Medin (Eds.), *Moral judgment and decision making. Psychology of learning and motivation* (Vol. 50, pp. 341-364). San Diego, CA: Academic Press.

Beauregard, K. S., & Dunning, D. (1998). Turning up the contrast: Self-enhancement motives prompt egocentric contrast effects in social judgments. *Journal of Personality and Social Psychology, 74*, 606-621.

Beauregard, K. S., & Dunning, D. (2001). Defining self worth: Trait self-esteem moderates the

use of self-serving trait definitions in social judgment. *Motivation and Emotion, 25*, 135-162.

Beck, L., McCauley, C., Segla, M., & Hershy, L. (1988). Individual differences in prototypicality judgments about trait categories. *Journal of Personality and Social Psychology, 55,* 286-292.

Blickle, G. (2003). Convergence of agents' and targets' reports on intraorganizational influence attempts. *European Journal of Psychological Assessment, 19*, 40-53.

Bogwald, K. P. (2001). Do patients and their therapists agree on the content of treatments? *The Journal of Nervous and Mental Disease, 189*, 830-837.

Bonanno, G. A. (2004). Loss, trauma, and human resilience. *American Psychologist, 59*, 20-28.

Bonanno, G. A., Field, N. P., Kovacevic, A., & Kaltman, S. (2002). Self-enhancement as a buffer against extreme adversity: Civil war in Bosnia and traumatic loss in the United States. *Personality and Social Psychology Bulletin, 28,* 184–196.

Bonanno, G. A., Rennicke, C., & Dekel, S. (2005). Self-Enhancement Among High-Exposure Survivors of the September 11th Terrorist Attack: Resilience or Social Maladjustment? *Journal of Personality and Social Psychology, 88*, 984-998.

Brunot, S. and Sanitioso, R.B. (2004). Motivational influence on the quality of memories: recall of general autobiographical memories related to desired attributes. *European Journal of Social Psychology, 34*, 627–635.

Camerer, C., & Lovallo, D. (1999). Overconfidence and excessive entry: An experimental approach. *American Economic Review, 89,* 306-318.

Cantor, N., Mischel, W., & Schwartz, J. C. (1982). A prototype analysis of psychological situations. *Cognitive Psychology, 14*, 45-77.

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.

Ditto, P. H., Pizarro, D. A., & Tannenbaum, D. (2009). Motivated moral reasoning. In B. H. Ross, D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin (Eds.), *Psychology of learning and motivation, Vol. 50: Moral judgment and decision making* (pp. 307-338). San Diego, CA: Academic Press.

Donaghue, N., & Smith, N. (2008). Not half bad: Self and others' judgments of body size and attractiveness across the life span. *Sex Roles, 58*, 875-882.

Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York: Psychology Press.

Dunning, D., & Beauregard, K. S. (2000). Regulating impressions of others to affirm images of the self. *Social Cognition, 18*, 198-222.

Dunning, D., & Cohen, G. L. (1992). Egocentric definitions of traits and abilities in social judgment. *Journal of Personality and Social Psychology, 63*, 341-355.

Dunning, D., & Hayes, A. F. (1996). Evidence for egocentric comparison in social judgment. *Journal of Personality and Social Psychology, 71*, 213-229.

Dunning, D., Heath, C., & Suls, J. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*, 69-106.

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12*, 83–87.

Dunning, D., Leuenberger, A., & Sherman, D. A. (1995). A new look at motivated inference: Are self-serving theories of success a product of motivational forces? *Journal of*

Personality and Social Psychology, 59, 58-68.

Dunning, D., & McElwee, R. O. (1995). Idiosyncratic trait definitions: Implications for self-description and social judgment. *Journal of Personality and Social Psychology, 68*, 936-946.

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*, 83-86.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology, 57*, 1082-1090.

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

Effron, D. A., Camerson, J. S., & Monin, B. (2009). Endorsing Obama licenses favoring whites. *Journal of Experimental Social Psychology, 45*, 590-593.

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware? Further explorations of (lack of) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes, 105*, 98-121.

Eidelman, S., & Biernat, M. (2007). Getting more from success: Standard raising as esteem maintenance. *Journal of Personality an Social Psychology, 92*, 759-774.

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.

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.

Fehr, B. (1988). Prototype analysis of the concepts of love and commitment. *Journal of Personality and Social Psychology, 55*, 557-579.

Fehr, B. (2004). Intimacy expectations in same-sex friendships: A prototype interaction-pattern model. *Journal of Personality and Social Psychology, 86*, 265-284.

Felson, R. B. (1981). Ambiguity and bias in the self-concept. *Social Psychology Quarterly, 44*, 64-69.

Freedman, D. G. (1974). *Human infancy: An evolutionary perspective.* Hillsdale, NJ: Lawrence Earlbaum Asssociates.

Gibson, B., & Oberlander, E. M. (2008). Wanting to appear smart: Hypercriticism as an indirect impression-management strategy. *Self and Identity, 7, 380-392.*

Goffman, E. (1955). On facework: An analysis of ritual elements in social interaction. *Psychiatry, 18*, 213-231.

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

Graham, J., Haidt, J., & Nosek, B. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*, 1029-1046.

Greve, W., & Wentura, D. (2003). Immunizing the self: Self-concept stabilization through reality-adaptive self-definitions. *Personality and Social Psychology Bulletin, 29*, 39-50.

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

Hayward, M. L. A., & Hambrick, D. C. (1997). Explaining the premiums paid for large acquisitions: Evidence of CEO hubris. *Administrative Science Quarterly, 42*, 103-127.

Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review, 106*, 766-794.

Horowitz, L. M., & Turan, B. (2008). Prototypes and personal templates: Collective wisdom and individual differences. *Psychological Review, 115*, 1054-1068.

Hsee, C. K. (1996). Elastic justification: How unjustifiable factors influence judgments. *Organizational Behavior and Human Decision Processes, 62*, 330-337.

Icheiser, G. (1943). Misinterpretations of personality in everyday life and the psychologist's frame of reference. *Character and Personality, 12*, 145–160.

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.

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly, 24*, 163-204.

Klein, W. M., & Kunda, Z. (1993). Maintaining self-serving social comparisons: Biased reconstruction of one's past behaviors. *Personality and Social Psychology Bulletin, 19*, 732-739.

Köhler, W. (1930). *Gestalt psychology*. London: Bell.

Kruger, J. M., & 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.

Kunda, Z., & Sanitioso, R. (1989). Motivated changes in the self-concept. *Journal of Experimental Social Psychology*, *25*, 272-285.

Kurman, J. (2003). The role of perceived specificity level of failure events in self-enhancement and in constructive self-criticism. *Personality and Social Psychology Bulletin*, *29*, 285-294.

Kurt, A. & Paulhus, D.L. (2008). Moderators of the adaptive value of self enhancement: Operationalization, domain, and form of adjustment. *Journal of Research in Personality, 42*, 839-853.

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.

Leslie, A. M., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect: Theory of mind and moral judgment. *Psychological Science, 17*, 421-427.

Lord, C. G., Desforges, D. M., Fein, S., Pugh, M. A., & Lepper, M. R. (1994). Typicality effects in attitudes toward social policies: A concept-mapping approach. *Journal of Personality and Social Psychology, 66*, 658-673.

Lord, C. G., & Lepper, M. R. (1999). Attitude representation theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 31, pp. 265-343). San Diego, CA: Academic Press.

Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *Journal of Finance , 60*, 2661-2700

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, *45*, 633-644.

McCrae, R. R., Stone, S. V., Fagan, P. J. & Costa, P. T. (1998). Identifying causes of disagreement between self-reports and spouse ratings of personality. *Journal of Personality, 66*, 285-313.

MacDonald, H. A., Sulsky, L. M., & Brown, D. J. (2008). Leadership and perceiver cognition: Examining the role of self-identity in implicit leadership theories. *Human Performance*, *21*, 333-353.

McElwee, R. O., Dunning, D., Tan, P. L., & Hollmann, S. (2001). Evaluating others: The role of who we are versus what we think traits mean. *Basic and Applied Social Psychology, 23*, 123-136.

Paulhus, D. L. (1998). Intrapersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology, 75*, 1197-1208.

Paulhus, D. L., & Landolt, M. A. (2000). Paragons of intelligence: Who gets nominated and why. *Canadian Journal of Behavioural Science*, *32*, 168-177.

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.

Robins, R. W., & John, O. P. (1997). Effects of visual perspective and narcissism on self-perception: Is seeing believing? *Psychological Science, 8*, 37-42.

Roese, N. J., & Olson, J. M. (2007). Better, stronger, faster: Self-serving judgment, affect regulation, and the optimal vigilence hypothesis. *Perspectives on Psychological Science, 2*, 124-141.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology, 104*, 192-233.

Rosch, E. (1978). Principles of categorization. In E. Rosch, & B. B. Lloyd (Eds.), *Cognition*

and categorization (pp. 27-48). Oxford, England: Erlbaum.

Ross, L., & Ward, A. (1996). Naive realism: Implications for social conflict and

misunderstanding. In T. Brown, E. Reed, and E. Turiel (Eds.),*Values and knowledge* (pp.

103-135). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rushton, J. P. (1999). Ethnic differences in temperament. In Y. –T. Lee, C. R. McCauley, & J.

G. Draguns (Eds.), *Personality and Person Perception Across Cultures* (pp. 45-63).

Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Sanitioso, R. B. (2004). Desired self and biased judgment of past behaviors. *Cahiers de

Psychologie Cognitive, 22*, 117-132.

Sanitioso, R.B., & Niedenthal, P.M. (2006). Motivated self-perception and perceived ease in

recall of autobiographical memories. *Self and Identity, 5*. 73-84.

Sanitioso, R. B., & Wlodarski, R, (2004). In search of information that confirms a desired self-

perception: Motivated processing of social feedback and choice of social interactions.

Personality and Social Psychology Bulletin, 30, 412-422.

Santos-Pinto, L., & Sobel, J. (2005). A model of positive self-image in subjective assessments.

American Economic Review, 95, 1386-1402.

Schneider, S. L. (2001). In search of realistic optimism: Meaning, knowledge, and warm

fuzziness. *American Psychologist, 56*, 250-263.

Sedikides, C., & Gaertner, L., & Toguchi, Y. (2003). Pancultural self-enhancement. *Journal of

Personality and Social Psychology, 84*, 60-79.

Solomon, S., Greenberg, J., & Pyszczynski, T. (1991). Terror management theory of self-

esteem. In C. R. Snyder & D. R. Forsyth (Eds.), *Handbook of social and clinical

psychology: The health perspective* (pp. 21-40). Elmsford, NY: Pergamon Press.

Story, A. L. (2003). Similarity of trait construal and consensus in interpersonal perception. *Journal of Experimental Social Psychology, 39*, 364-370.

Suls, J., Lemos, K., & Stewart, H. L. (2002). Self-esteem, construal, and comparisons with the self, friends, and peers. *Journal of Personality and Social Psychology, 82*, 252-261.

Swann, W. B., Jr. (1983). Self-verification: Bringing social reality into harmony with the self. In J. Suls & A. G. Greenwald (Eds.), *Social psychological perspectives on the self* (Vol. 2, pp. 33-66). Hillsdale, NJ: Erlbaum.

Tannenbaum, D., Ditto, P. H., & Pizarro, D. A. *Different moral values produce different judgments of intentional action*. Unpublished manuscript, 2008.

Taris, T. W. (1999). Describing behaviors of self and others: Self-enhancing beliefs and language abstraction level. *European Journal of Social Psychology, 29*, 391-396.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.

Taylor, S. E., & Brown, J. D. (1994). Positive illusions and well-being revisited: Separatiing fact from fiction. *Psychological Bulletin, 116*, 21-27.

Taylor, S. E., & Gollwitzer, P. M. (1995). Effects of mindset on positive illusions. *Journal of Personaity and Social Psychology, 69*, 213-226.

Taylor, S. E., Lerner, J. S., Sherman, D. K., Sage, R. M., & McDowell, N. K. (2003). Are self-enhancing cognitions associated with healthy or unhealthy biological profiles? *Journal of Personality and Social Psychology, 85*, 605-615.

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology,90*, 288-307.

Tsai, J. L., Louie, J. Y., & Uchida, Y. (2007). Learning what feelings to desire:

Socialization of ideal affect through children's storybooks. *Personality and Social Psychology Bulletin, 33*, 17-30.

Tsai, J. L., Miao, F. F., & Seppala, E. (2007). Good feelings in Christianity and Buddhism:

Religious differences in ideal affect. *Personality and Social Psychology Bulletin, 33*, 409-421.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 58,* 806-820.

Wentura, D., & Greve, W. (2004). Who wants to be…erudite? Everyone! Evidence for automatic adaptation of trait definitions. *Social Cognition 22*, 30-53.