Construct Accessibility and Interpretation of Self-Behaviors: Tracing and Reducing the Signatures of the Self-Protection Motive and the Self-Enhancement Motive

Constantine Sedikides
University of Southampton

John J. Skowronski
Northern Illinois University

WORD COUNT (text): 6535

Corresponding Author:
Constantine Sedikides, Center for Research on Self and Identity, School of Psychology, University of Southampton, Southampton, England; cs2@soton.ac.
Self-evaluation is not neutral. Other-evaluation can be, but self-evaluation rarely is. Instead, self-evaluation is guided by motives, of which the most prominent are self-protection and self-enhancement. Self-protection refers to avoiding, minimizing, misinterpreting, or discarding information that has unfavorable implications for the self. Self-enhancement, on the other hand, refers to pursuing, magnifying, overinterpreting, or fully endorsing information that has favorable implications for the self. (For definitional nuances, see: Alicke & Sedikides, 2009; Sedikides & Gregg, 2003, 2008; Sedikides & Strube, 1997.)

**Potency and Prevalence of Self-Enhancement Strivings and Self-Protection Strivings**

The self-protection motive and the self-enhancement motive manifest themselves through a seemingly endless repertoire of cognitions, emotions, and behaviors (Alicke & Sedikides, 2011). After a comprehensive review of the literature, Hepper, Gramzow, and Sedikides (2010) identified 60 major self-enhancement/self-protection strivings. These authors then created a questionnaire that represented the strivings and asked participants to judge how characteristic or typical each striving was of them. For example, to operationalize the better-than-average striving (Alicke & Govorun, 2005), participants imagined “thinking of yourself as generally possessing positive traits or abilities to a greater extent than most people do” and subsequently rated how characteristic this striving was of them.

Hepper et al. (2010) used exploratory and confirmatory factor analyses to distill the 60 strivings into four “families.” The first one was termed *positivity embracement*. It consisted of 10 strivings that pertained to the acquisition or retention of positive (i.e., self-enhancing) feedback or the maximization of expected success. Examples are: making self-serving attributions for success (Mezulis, Abramson, Hyde, & Hankin, 2004), engaging in self-promoting social interactions (Leary, 2004), and remembering...
positive feedback better than negative feedback (Skowronski, 2011). The second family was termed favorable construals. It consisted of six strivings that pertained to forming flattering construals of the self in the social world. Examples are: positive illusions (Taylor & Brown, 1988), comparative optimism (Weinstein, 1980), and self-favoring interpretations of ambiguous or negative feedback (Critcher, Helzer, & Dunning, 2011). The third family was termed defensiveness. It consisted of 18 strivings that pertained to the protection of the self from threat. Examples are: self-handicapping (Zuckerman & Tsai, 2005), defensive pessimism (Norem & Cantor, 1986), outgroup derogation (Fein & Spencer, 1997), moral hypocrisy (Batson & Collins, 2011), and self-serving attributions for failure (Campbell & Sedikides, 1999). The fourth and final family was termed self-affirming reflections. It consisted of six strategies that pertained to securing positive self-views or outcomes when faced with the potential for negative outcomes. Examples are: downward counterfactual thinking (Sanna, Chang, & Meier, 2001), temporal comparison (Wilson & Ross, 2001), and focusing on strengths, values, or relationships (Sherman & Cohen, 2006). Importantly, these four families of striving were validated not only in Western cultures (Hepper et al.) but also in an East-Asian culture (i.e., China; Hepper, Sedikides, & Cai, 2011).

**Are Self-Enhancement and Self-Protection Strivings Beneficial?**

In all, the data suggests that the self-enhancement and self-protection motives are potent and prevalent. More importantly, the motives confer psychological and pragmatic benefits to the individual. These include better psychological health (Sedikides, Gregg, & Hart, 2007), better social adjustment (Donnellan, Trzesniewski, Robins, Moffitt, & Caspi, 2005), improved physical health (Taylor, Lerner, Sherman, Sage, & McDowell, 2003), more effective coping with traumatic life events (Bonanno, Rennicke, & Dekel, 2005), greater persistence in the face of adversity (Taylor & Brown, 1988), and
advancement of one’s interests or goals (Alicke & Sedikides, 2009).

Nonetheless, the two motives are not free of personal, interpersonal, and behavioral liabilities. Personal liabilities include ill-considered risk taking (Baumeister, Heatherton, & Tice, 1993), imprudent action planning (Oettingen & Gollwitzer, 2001), and—perhaps as a result—increased likelihood of disengagement from academic studies (Robins & Beer, 2001). Interpersonal liabilities include being perceived and treated unfavorably by others. For example, following a brief infatuation period, peers come to consider habitual self-enhancers as defensive, condescending, and hostile (Paulhus, 1988), and tend to deride or socially isolate them (Schlenker & Leary, 1982). Finally, behavioral liabilities include actions that lead to illness, injury, and death. For example, individuals from temperate climates often sunbathe extensively, accepting the risks of sunstroke, sunburn, and skin cancer for the sake of feeling good and looking good among their peers (Leary, Tchividjian, & Kraxberger, 1994).

Are Self-Enhancement and Self-Protection Strivings Controllable?

Given the alarming set of potential costs associated with self-enhancement and self-protection, researchers have asked whether such strivings are controllable. Fortunately, empirical findings suggest that both interpersonal factors and intrapersonal factors place boundaries on the scope of self-enhancement and self-protection effects.

Interpersonal Limits

One set of limiting factors reflect interpersonal contexts. Examples of these are the relational and social context.

Relational context. Relationship closeness restrains self-enhancement strivings. For example, in a study by Campbell, Sedikides, Reeder, and Elliot (2000; see also McCall, Reno, Jalbert, & West, 2000), some participants worked collaboratively with a stranger, whereas others worked with a friend, on an interdependent-outcomes task billed
as a creativity test. In this task, participants generated as many unique uses as possible for several objects (e.g., candle, brick). Then, participants received either success or failure feedback at the dyadic level (i.e., based on the combined score). Strangers (or distant participants) displayed the self-serving bias: They blamed the partner for dyadic failure, and claimed personal credit for dyadic success. However, friends (or close participants) refrained from the self-serving bias: They shared responsibility for both dyadic failure and dyadic success.

Notably, this pattern does not depend on expectations for future and rewarding interactions with the close partner or on expectations for relationship maintenance. Stated otherwise, relationship closeness per se suffices for a reduction in self-enhancement. Data supporting this claim were provided by Sedikides, Campbell, Reeder, and Elliot (1998). These authors tested only unacquainted participants who did not anticipate interacting with one another and also promised not to discuss the experiment in incidental (albeit unlikely) encounters on a large academic campus. Closeness was induced experimentally in half of participants through a reciprocal and escalating self-disclosure procedure (i.e., taking turns in asking questions that required increasingly intimate answers). Again, distant participants displayed the self-serving bias, but close participants did not.

**Social context.** The social context also places constraints on the scope of self-enhancement and self-protection effects. Verifiability is a case in point: People self-enhance less on attributes that are easy rather than difficult to verify. For example, athletes lionize themselves less on unambiguous (e.g., speed, size, ball heading) than ambiguous (e.g., mental toughness, coordination, soccer ability) attributes (Felson, 1981; Van Yperen, 1992); students glorify their performance less on concrete than global domains (Kurman & Eshel, 1998; Willard & Gramzow, 2009); job applicants exaggerate
their resumes less when they know that the organization can confirm the information (Donovan, Dwight, & Hurtz, 2003); and people describe themselves less positively on traits that are easy (e.g., intelligent) than (e.g., fair) difficult to corroborate (Allison, Messick, & Goethals, 1989; Van Lange & Sedikides, 1998).

One reason that verifiability constrains self-enhancement and self-protection is because of its potential for accountability. The importance of accountability is illustrated in research by Sedikides, Herbst, Hardin, and Dardis (2002, Experiment 1). In this research, participants wrote an opinion essay (e.g., “Should the United States pursue exploration of the planet Saturn?”) and prepared to grade it. In the meantime, some participants were led to believe that they were accountable (i.e., had to explain, defend, and justify their grade) to another person, whereas other participants were led to believe that they were unaccountable. Essay grading followed. Accountable participants gave their essays lower grades compared to unaccountable ones. Furthermore, accountable participants who were identifiable to an evaluative audience were particularly likely to assign their essays lower grades (Sedikides et al., Experiments 2-3).

Intrapersonal Limits

Importantly, not only interpersonal factors, but also intrapersonal factors bound self-enhancement and self-protection. Examples of such factors are self-focus and introspection.

Self-focus. An internal focus reduces self-enhancement. For example, the presence of a mirror decreases inflation of reported SAT scores (Pryor, Gibbons, Wicklund, Fazio, & Hood, 1997). Relatedly, accountable (and thus less self-enhancing) participants focus on their weaknesses as essay writers (Sedikides & Herbst, 2002). Moreover, in open-thought protocols, accountable participants express doubts about their competence as essay writers, reflect on the grueling experience of writing essays in the
past, wonder how poor many of their past essays must have been, and recollect how critical other people were of their essays (Sedikides et al., 2002, Experiment 4). Attentional focus on weaknesses, then, is tantamount to self-criticism. Indeed, the presence of a mirror during a writing task fosters self-criticism (Heine, Takemoto, Moshalenko, Lasaleta, & Henrich, 2008). A reason for self-enhancement curtailment under self-focus is that self-focus draws attention toward one’s inner standards, and thus highlights the discrepancy between ideal self and the actual self (Silvia & Duval, 2001).

**Introspection.** Introspection is a special case of self-focus and also curbs self-enhancement and self-protection effects. Support for this idea was provided by Sedikides, Horton, and Gregg (2007, Experiments 1-2). Participants thought carefully (i.e., introspected) about the reasons why they possessed or did not possess particular traits (e.g., kind, honest, trustworthy or unkind, dishonest, untrustworthy), listed these reasons, and then rated themselves on these traits. Participants who introspected about positive traits rated themselves as less possessing of such traits compared to a non-introspecting control group; thus, these persons experienced a drop in self-enhancement. However, participants who introspected about negative traits rated themselves as more possessing of such traits compared to a control group; thus, these persons experienced a drop in self-protection. These decreases in self-enhancement and self-protection were due to accompanying reductions in self-certainty (Sedikides et al., Experiment 3; see also Petty, Brinol, & Tormala, 2002). Moreover, it is possible that the accessibility of trait constructs (Sedikides & Skowronski, 1991) may have played a role in the observed effects. For example, introspection might have diminished the endorsement of positive traits (e.g., kind) by increasing the accessibility of negative traits (e.g., unkind). Regardless of the exact reasons, self-enhancement and self-protection motives seem to be relatively impotent when individuals introspect.
Flexibility in Self-Thought: The Influence of Construct Accessibility

On a general level, the ideas expressed in the paragraphs above suggest that peoples’ thoughts about themselves can exhibit considerable variation across time and context. Indeed, considerable research documents the malleability and flexibility of self-conceptions (Fazio, Effrein, & Falender, 1981; Markus & Kunda, 1986; Schwarz, Bless, Strack, Klumpp, Rittenauer-Schatka, & Simmons, 1991), reflecting the fact that self-perceptions manifest both temporal inconsistency and cross-situational flexibility (DeSteno & Salovey, 1997).

Researchers have undertaken the exploration of factors that influence self-perception variation. One such factor is construct accessibility. Constructs vary in their accessibility (roughly, their “state of readiness to be used”), and this variability influences later cognitive processing about others (DeCoster & Claypool, 2004; Sedikides & Skowronsks, 1991). This perspective can also be applied to self-perceptions, suggesting that variability in the accessibility of self-related traits can account for short-term variations in judgments about the self (Schubert & Häfner, 2003; Smeesters, Wheeler, & Stapel & Koomen, 2001; Kay, 2010; Wyer, Calvini, & Nash, 2010). For example, parallel distributed processing conceptions of self-representations suggest that self-concepts are recomputed each time a judgment is required, and that such recomputations are influenced by those portions of the self-evidentiary base that are made accessible by situationally-triggered constructs (Van Overwalle & Labiouse, 2004).

Existing data already document some of the roles that construct accessibility plays in self-thought. For example, Chatard, Guimond, and Selimbegovic (2007, Study 2) primed gender stereotypes by asking student participants to evaluate stereotype-consistent statements (e.g., “Men are gifted in mathematics,” “Women are gifted in arts”) and then asked the participants to report their grades in maths and arts courses. Male
participants reported exaggerated math grades when gender stereotypes were accessible, whereas female participants reported exaggerated arts grades when arts stereotypes were accessible. Similarly, Marx and Stapel (2006) showed that both female participants and male participants primed with the stereotype of men as emotionally insensitive reported decreased emotional sensitivity. Finally, non-African Americans primed with the African-American stereotype reported both feeling especially close to African-Americans and heightened feelings of aggression (DeMarree, Wheeler, & Petty, 2005; Galinsky, Wang, & Ku, 2008).

**Variation in Behavior-Based Inferences about the Self**

Such variation across time and circumstance can apply to the interpretations that people give to their behaviors. Indeed, people sometimes do not know why they do what they do before they do it. Instead, they may infer a behavior’s meaning and engage in consequent self-inferences after having enacted the behavior (Stone & Cooper, 2001). Clearly, attempts at attitude change via induced compliance use this idea (Burger & Caldwell, 2003). In an example of induced compliance process, consider that a homeowner might be induced without obvious pressure or incentive to place a sign endorsing a politician in the homeowner’s front yard. Later, the homeowner may try to deduce the motivation underlying the behavior. He may recognize that placing a sign in one’s yard can mean that one supports the politician. Hence, he may infer that he is a supporter. Another interpretation that might be made by the homeowner is that he simply wanted to assist someone in need of help. Yet another possible interpretation that the homeowner might give to the behavior is that he wanted to avoid unpleasant interpersonal confrontation and said “yes” simply to make the requesters go away.

Not only can such variation apply to inferences about motives, it can also apply to inferences about traits and dispositions. Consider the hypothetical example of Laura, a
golf lover. While playing golf, Laura might perceive that she is hitting the ball well on the practice tee, and think that she might finally be improving at the game. When she subsequently holes a chip shot on the first hole, she may see this as confirmation that her golf is improving. However, on another occasion, Laura might be in the midst of a swing change and might be looking at herself as someone who is currently unskilled at golf. Thus, when she subsequently holes a chip shot on the first hole, she might conclude that the shot was a very lucky stroke, indeed.

Curiously, although past literature has established that priming can modify self-judgments, until recently research had not shown that priming could directly alter interpretation of self-behavior (for related ideas, see Sedikides & Herbst, 2002; Sedikides et al., 2007; Silvia & Duval, 2001). This was an issue that our team addressed in a recent set of studies (Skowronski, Sedikides, Heider, Wood, & Scherer, 2010).

**Self-Enhancement and Self-Protection Motivation as Moderators of Priming Effects**

Those studies also addressed a second question: To what extent are priming effects moderated by self-enhancement or self-protection motivation? To illustrate this issue, let us return to our example of the golfer, Laura. Might it be easier to induce Laura to see her performance as a result of her increasing skill (a self-serving attribution) than luck? The notions of self-enhancement and self-protection suggest that this might be so. That is, if people are motivated to view themselves positively, it might be especially easy to use priming to induce Laura to interpret a positive performance as reflecting her skill at a task than as reflecting luck (a self-enhancement effect); it might also be easier to induce Laura to interpret a negative performance as reflecting bad fortune instead of reflecting a deficiency in her skill level (a self-protection effect).

The Skowronski et al. (2010) experiments took a step toward addressing such issues by examining the extent to which priming effects were observed for the self (or
not) and comparing the extent to which these effects were observed for a hypothetical other. To understand the logic underlying the studies, consider a participant being exposed to a priming manipulation that heightens the accessibility of the trait hostile. Then, imagine that the participant reads a story in which a protagonist behaves in an ambiguous manner that could be construed as hostile.

When the story protagonist is not the same person as the participant, participants rate the story’s protagonist as especially hostile (Srull & Wyer, 1979). However, what would happen if participants were to imagine that they were the story protagonist? This is where things get theoretically interesting. Extrapolating from the findings in the other-perception literature, one possibility is that the priming manipulation would affect interpretations of self-behavior in the same way as it affects interpretations of the same behavior when performed by another person. That is, those construals would be consistent with the implication of the primed construct and would influence subsequent self-judgments in a trait-congruent manner. Thus, when the construct of meanness is made accessible, participants might be more likely to interpret an ambiguous self-behavior (e.g., telling a friend that her hairdo is ugly) as reflecting meanness rather than honesty (a viable alternative interpretation). Accordingly, given the performance of a recent mean behavior, later self-judgments would reflect heightened meanness.

However, the notions of self-enhancement and self-protection suggest that negative self-labeling might be especially difficult. That is, given the potency and prevalence of self-protection and self-enhancement motivation, the possibility that construct accessibility will influence negative behavior interpretation and negative self-inference is rather limited. Indirect support for this notion comes from research showing that an experimentally-induced focus on the trait implications of negative self-behavior yields maximal attitude change, presumably as a consequence of heightened arousal that
accompanies such a focus (Blanton, Cooper, Skurnik, & Aronson, 1997). A second empirical justification for this hypothesis comes from literature that compares other-perception to self-perception. This literature indicates that perceptions of these two classes of actors often differ, even when those perceptions are derived from the same objective data (Pronin, Lin, & Ross, 2002; Sedikides & Green, 2009).

Hence, taken to an extreme, the self-enhancement/self-protection viewpoint posits that the meanness primes should not have an assimilative effect on the interpretation of self behavior, despite having such an effect on the interpretation of the behavior of another. This viewpoint would also be supported, if the priming manipulation produced weaker assimilative effects for the self than other.

**Construct Accessibility and Interpretation of Self-Behavior:**

**The Initial Meanness Priming Studies**

Our research team tested these ideas in experiments that used a priming manipulation to vary the accessibility level of the meanness construct (Skowronski et al., 2010, Experiment 2; see also Experiment 1). Replicating methods used in past research (Sedikides & Skowronski, 1991), a sentence unscrambling task was used as the manipulation. Participants had 8 min to rearrange 60 scrambled word sequences (between 4-6 words each) into a grammatical sentence that comprised at least three words. There were two experimental conditions. In the meanness priming condition, 50 of the 60 word sequences, when unscrambled, implied meanness. For example, “cry them let make” could be rearranged to “make them cry” or “let them cry,” and “cat the kick his” could be rearranged to “kick the cat” or “kick his cat.” In the neutral priming condition, none of the word sequences reflected meanness when unscrambled.

After completing a filler task, participants received a packet containing a vignette (based on Srull & Wyer, 1979) that described the behaviors of an actor. Our objective
was to equate other-perception with self-perception in order to be able to compare the two conditions. For one half of participants the actor was another person (Terry), whereas for the other half the actor was the self. Participants in the Terry condition imagined that they were about to meet a person named Terry and that the vignette described Terry’s recent day with a friend. Participants were instructed to consider the vignette’s implications for Terry’s personality. Participants in the self condition imagined that they were about to meet another person who wanted to know in advance what the participant was like. Participants imagined that, in response to this request, they had generated the vignette as a description of a recent day they spent with a friend. (The vignette featured the pronouns I, me, or my). Participants were instructed to think about personality characteristics they might have based on the vignette alone.

The behaviors included in the vignette were ambiguously mean. Each behavior was accompanied by qualifying situational information intended to increase further the ambiguity of its trait implications. Specifically, the behaviors (and the qualifiers) were: (1) did not buy anything from a salesman (who knocked loudly); (2) said that they would not pay the rent until the apartment was painted (which was supposed to be pained two years ago); (3) told the mechanic that the car needed to be fixed this week, not next week (needed the car for his job); (4) asked the clerk about getting money back for a gadget (which had faulty electronics); and (5) yelled at the referee who made a bad call at a basketball game (call was overruled by another referee).

After reading the vignette at their own pace, participants judged the personality of the target (Terry or the self, depending on condition). The ratings reflected meanness-related traits (e.g., aggressive, hostile, ill-tempered, rude) and meanness-unrelated or control traits (e.g., creative, spiritual, disorganized, undependable).

The priming manipulation was effective. Collapsing across targets, participants in
the meanness priming condition rated the target as meaner than those in the neutral condition. Furthermore, participants in the two conditions did not differ in their target ratings on the control traits, so this effect was limited to the meanness ratings. Importantly, evidence also emerged for the operation of the self-protection/self-enhancement motives. Participants rated themselves more positively (on both meanness-related and control traits) than they rated Terry.

Of particular relevance, though, was the interaction between priming condition and target. If self-protection/self-enhancement motivations intervene in the interpretation of self-behaviors, then meanness accessibility effects should have been strong in other-perception but weak (or non-existent) in self-perception. This did not occur: The construct accessibility effects that emerged were virtually identical for Terry and self. In other words, priming the construct of meanness exacerbated meanness judgments of an ambiguously-behaving actor, regardless of whether this actor was Terry or the self. This outcome is not congenial to the self-protection/self-enhancement viewpoint.

The Pursuit of Moderation Effects in the Self-Priming Paradigm: Introducing Self-Threat

In reviewing these results, Skowronski et al. (2010) considered the possibility that failure to find support for the moderation by target may have been due to the fact that the experimental paradigm did not strongly prompt the self-protection motive. This motive is instantiated by threat (Alicke & Sedikides, 2009; Campbell & Sedikides, 1999; Sedikides & Green, 2009), and participants may not have felt especially threatened by their negative interpretations of their behaviors or by giving themselves a negative trait rating based on these behaviors.

In Skowronski et al. (2010, Experiment 3), our research team pursued the idea that self-protection motivation, induced by self-threat, would decrease the likelihood that
participants would rate themselves as mean. In the experiment, half of participants completed a judgment task intended to threaten the self. The other half of participants (i.e., the control group) completed a judgment task intended to enhance the self. In addition, we orthogonally manipulated target: Two additional conditions exposed participants to judgment tasks in which other persons were either threatened or enhanced. Inclusion of these conditions attempted to ensure that, if self-threat effects emerged, they could specifically be attributed to self-threat instead of to general threat.

Let us be more specific about this task. In the \textit{self} condition, participants considered their own chances of experiencing an event. Examples are: (1) “Some experts have claimed that 8 out of every 10 college students will experience mental deficits in old age because of their drug use in college. What do you think your chances are of experiencing mental deficits in old age because of your drug use in college?”; (2) “Some experts have claimed that high levels of sexual activity in college will lead to 4.5 out of every 10 college students contracting a sexually-transmitted disease. What do you think your chances are of contracting a sexually-transmitted disease?”; and (3) “Some experts have claimed that because of widespread gun availability, 1 out of every 6 people will be the victim of a shooting. What do you think your chances are of being the victim of a shooting?” In the \textit{other} condition, participants considered other peoples’ chances of experiencing the same events. In addition, the events varied by valence across groups. In the \textit{threat} condition, participants considered the occurrence of negative life events (for self or others). However, in the \textit{enhancement} condition, participants considered the occurrence of positive life events (for self or others). Participants responded to each event by providing a numerical probability estimate ranging from 0% to 100%.

Subsequently, participants engaged in the priming task, which was identical to that used in the experiment described earlier in this chapter. As before, half of
participants were exposed to the *meanness prime* condition (i.e., 60 word sequences relevant to meanness); half were exposed to the *neutral* condition (i.e., no word sequences relevant to meanness). Participants then read a meanness-ambiguous paragraph about either Terry or the self (as in the previous experiment) and rendered judgments on several sets of traits. The first set contained the meanness-related traits used in the prior experiment. The second set pertained to assertiveness (e.g., assertive, firm, resolute). The third set contained positive traits that were unrelated to meanness (e.g., creative, fashionable, logical), and the fourth set contained negative traits that were unrelated to meanness (e.g., dull, lazy, shallow). We included the latter three sets of traits in order to explore the specificity of the effects obtained in the initial studies.

As in Experiment 1, results suggested that the priming manipulation was effective. Participants in the meanness priming (relative to neutral) condition perceived the target to be especially mean; this effect was specific to meanness and did not emerge on the other three sets of control traits. The event probability task manipulation was effective, and also showed evidence of a self-protective bias. Participants who received the threatening (compared to the flattering) version of the task assigned a lower probability to event occurrences. Moreover, in comparison to the ratings provided when the questions asked about other people, participants assigned lower occurrence probabilities to threatening events when the target was the self, and correspondingly assigned higher occurrence probabilities to flattering events when the target was the self.

Despite such evidence, the meanness trait ratings were not moderated by the other manipulations. Instead, the results showed that priming the meanness construct had assimilative effects on meanness judgments, and this pattern occurred equally for Terry and the self. Notably, this lack of moderation by target occurred despite evidence elsewhere in the trait ratings suggesting the presence of self-enhancement strivings:
Participants judged the self as less mean, and as more positive on the control traits, than Terry. Moreover, the self-threat induction manipulation was similarly impotent: It did not moderate the effect of meanness construct accessibility on self-judgments of meanness, despite the fact that evidence of exposure to the self-threat manipulation could be seen elsewhere in the trait judgments.

**Deliberative Versus Direct Construal of Self-Behavior: Different Routes to Self-Judgment?**

The empirical findings that we discussed so far (Skowronski et al., 2010, Experiment 1-3) showed that variations in construct accessibility can influence interpretation of one’s behavior and subsequent self-judgment in an assimilative manner, even when the accessible construct is negative (i.e., meanness). These findings suggest that, in the context of priming, self-enhancement effects are nulled when interpretation of self-behavior is involved. The findings are consistent with some past research on the topic. For example, when individuals self-focus while comparing their actual versus ideal performance standards (Silvia, 2001), when they focus on and analyze their weakness (Sedikides & Herbst, 2002), and when they explain introspectively why or why not they might possess certain traits (Sedikides et al., 2007), then they tone down the positivity of their self-inferences. Common to these studies is that all involve the considered and deliberative interpretation of self-behaviors. This same deliberative processing likely characterized the Skowronski et al. (2010) studies that we have so far described, and it caused us to wonder if this state of affairs was responsible for the absence of evidence favoring self-enhancement/self-protection motivation in the studies that we have described so far.

**Construct Accessibility and Construal of Self-Behavior: An Induced Compliance Experiment**
This reasoning led our team to conduct an additional experiment (Skowronski et al., 2010, Experiment 4). The methods used in this experiment differed from the methods used in our first three experiments. The idea underlying the experiment was to get participants to engage in a behavior (or to watch others engage in the behavior) that could be interpreted as either a reflection of the trait helpfulness or the trait dishonesty. This behavior (or observation) was executed after exposure to a priming manipulation that activated either the helpfulness or dishonesty trait. Importantly, participants were not aware of the true intent of the experiment as they were proceeding through these tasks. Hence, in contrast to Experiments 1-3, in Experiment 4 participants were not deliberatively processing their behaviors as they read about them. Instead, in the self condition, they engaged in these behaviors thinking that they were being performed for another purpose. It was only later that participants in the self condition were asked to make self-ratings into which those behaviors could be incorporated.

The subliminal priming task. To help avoid deliberative processing, a subliminal priming manipulation (based on a manipulation used by Farc, Crouch, Skowronski, & Milner, 2008) was used for activating trait constructs. At the start of the experiment, participants learned that they would take part in a brief study assessing their attentional abilities. Using parafoveal presentations of the priming stimuli, one third of participants encountered multiple primes intended to activate the construct dishonest (e.g., cheater, deceitful, hypocrite, liar). Another third encountered multiple primes intended to activate the construct helpful (e.g., aid, assist, generous, giving). The final third of participants encountered multiple neutral primes (e.g., something, between, said, there). An ensuing recognition memory test determined that participants were unaware of the priming stimuli.

The task: telling lies. Past literature has shown that lying is susceptible to self-
protection/self-enhancement. For example, at the conclusion of a diary study on lying in daily life, both university students and community adult volunteers estimated that they lied less frequently than other participants in the study (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996).

Accordingly, in Experiment 4 we used a lie-telling paradigm. The beauty of the paradigm is that it could be construed as reflecting the trait of helpful (e.g., because the lies were requested) or as a reflection of dishonesty (e.g., because the participant was lying). We were especially interested in whether the priming manipulation affected participants’ post-lie self-perceptions, and whether those self-perceptions were seemingly moderated by self-perception/self-enhancement motivations.

To test this latter idea, we implemented both a self condition and an observer condition. In the self condition, we used a variant of the induced compliance paradigm (Burger & Caldwell, 2003). We told participants that we would appreciate it, if they provided us with some fictitious stories for the experiment, for these were especially hard to obtain. As in the usual induced compliance procedure, we refrained from overly pressuring participants and tried as hard as we could to allow participants to maintain the illusion of free choice.

All participants agreed with the request. Next, they were given 10 min to compose five brief fictional autobiographical stories which they would tell while being recorded by a video camera. Having composed the stories, participants were seated in front of the video camera and conveyed each story, one at a time. They were instructed to adopt a conversational style, as if they were chatting to a friend or family member.

Finally, under the ruse of providing information so that experimenters could better decipher the stories, participants responded to questions about personality traits that might be characteristic of them. Some of the responses related to the trait dishonest
(e.g., dishonest, hypocritical, insincere, treacherous), others to the trait *helpful* (e.g., altruistic, charitable, compassionate, helpful), and others were neutral.

In the *observer condition*, after exposure to one of the priming manipulations, participants learned that they would view videos of other students telling fictional stories about themselves. They were given the instructions that participants in the self condition had received, and so they knew that participants provided the stories at the experimenter’s request. Subsequently, observers viewed one of the videos supplied by participants in the self condition. Finally, observers rated the story-tellers on the same traits as the ones on which participants had rated themselves.

**The empirical verdict.** The subliminal priming manipulation was effective. Participants who were primed with the construct *dishonest* (vs. helpful or neutral) rated themselves higher on dishonesty traits than on either helpful or neutral traits, and so did observers. Likewise, participants who were primed with the construct *helpful* (vs. dishonest or neutral) rated themselves higher on helpfulness traits than on either dishonesty or neutral traits, and so did observers. Furthermore, as in Experiments 1-3, the self-protection/self-enhancement motives were instantiated. Participants judged themselves as less dishonest than observers did, and they judged themselves as more helpful than observers did.

The key question in Experiment 4, however, hinged on whether the ratings of honesty and helpfulness showed evidence of moderation by the motives. Importantly, and in contrast to Experiments 1-3, Experiment 4 revealed these moderation effects. In comparison to the ratings provided in the other two priming conditions, participants primed with the trait *dishonest* perceived themselves as less dishonest than did the observers. Likewise, in comparison to the ratings provided in the other two priming conditions, participants primed with the trait *helpful* perceived themselves as more
helpful than did the observers.

**Can the results of Experiment 4 be attributed to the real nature of the behavior?** We submit that this empirical discrepancy reflects the extent to which participants engaged in deliberative processing of their behaviors. In our view, participants were likely engaging in deliberative processing in Experiments 1-3, but not in Experiment 4.

However, there were other differences between the experiment sets, and these differences may account for the discrepant results. For example, self behaviors were hypothetical in Experiments 1-3, but were real in Experiment 4. Might self-involvement be more potent in the latter than former case?

We do not believe that this methodological difference can explain the findings. Self-protection/self-enhancement effects have been observed in a variety of self-other comparison experiments that have implemented imaginary scenarios (Green, Sedikides, & Gregg, 2008; Sedikides & Green, 2004), and the effects of such comparisons have been equally potent across imaginary and real situations (Green, Sedikides, Pinter, & Van Tongeren, 2009; Sedikides & Green, 2000). Moreover, self-protection/self-enhancement effects were still evident in the supposedly uninvolving scenario procedures (Experiments 1-3), but they reflected statistically self/other main effects rather than the interaction between target and prime.

Hence, we consider it unlikely that the reality of the situation used in Experiment 4 magnified the need to self-protect or to self-enhance. Instead, we maintain that the experimental situation worked to bypass deliberative processing of self-behaviors. This allowed the action of the self-enhancing/self-protective motives to become manifest in interpretations of the prime-relevant behaviors, and these effects in turn became manifest in self-perceptions.
Concluding Notes

The research described in this chapter brings together literatures on other-perception and self-perception, construct accessibility, and self-evaluation. In pursuing this fusion, the research also provides another demonstration of the yin and yen of self-enhancement/self-protection processes.

One way it does so is to illuminate limitations on the actions of self-enhancement and self-protective strivings. Such strivings have many psychological and pragmatic benefits, but they also entail liabilities (Alicke & Sedikides, 2009; Sedikides et al., 2007). As such, research has looked into ways that such strivings can be curtailed. Some of these ways are interpersonal (e.g., relational context, social context), and some intrapersonal (e.g., self-focus, introspection).

One intrapersonal method that limits the scope of such strivings concerns the extent to which priming affects self-perceptions. A task entailing the deliberative processing of behaviors curtailed the positivity of self-inferences derived from priming-altered perceptions of behavior. Individuals primed with a negative trait, then exposed to behaviors that might reflect the trait, perceived themselves as negatively as they perceive another person. However, under conditions in which deliberative consideration of behavior is bypassed, as when subliminal primes are used in an attempt to influence the interpretation of real behavior in an induced compliance paradigm, then the action of the self-enhancement/self-protection motives do moderate the impact of the primes on behavior interpretation. Here, individuals primed with a negative trait perceived themselves less negatively than they were perceived by an observer, and those primed with a positive trait perceived themselves more positively than they were perceived by an observer.

Our findings open up interesting possibilities. For example, would priming of the
construct dishonest alter the outcome of the classic Festinger and Carlsmith (1959) insufficient justification study? Here, participants are subtly induced to behave in a dishonest manner (i.e., lying to another individual about a task’s interest value). Participants who lie for insufficient justification regard the task as more interesting. It is possible that dishonesty priming would cause participants to be less likely to manifest this attitudinal shift. Such participants may be more likely to interpret their behavior in dispositional terms (e.g., “I am dishonest”), an interpretation that would eliminate the need to justify the behavior through attitude change.

One also wonders about the ultimate behavioral implications of our findings. Certainly, there is some evidence that construct accessibility can affect self-behavior. For example, Marx and Stapel (2006) demonstrated that primes designed to activate gender stereotypes sometimes reduced mathematics tests performance in some females. Similarly, in DeMarree et al. (2005, Study 3), some participants exposed to a professor stereotype were especially influenced by a persuasive message; this influence occurred because the stereotype prompted an especially deep consideration of message arguments. Most relevantly, Hansen and Wänke (2009) primed participants with stereotypes of professors or stereotypes of cleaning ladies. Next, participants completed trivial pursuit questions. Those primed with the professor stereotype correctly answered more questions than those primed with the cleaning lady stereotype. These effects were mediated by a sense of self-efficacy, a finding that attests to the important role of self-construals.

Given these findings, it does not take much to imagine how self-perceptions related to dishonesty and helpfulness might be similarly influenced by the self-construals of ambiguous events. If I told a lie and concluded that I was dishonest, I might be especially likely to tell another lie. However, if I told a lie and concluded that I was
helpful, I might not increase my lie-telling tendencies (except, perhaps, if doing so could again be construed as a helpful act). From this perspective, then, the self-protection/self-enhancement motives might be particularly important to social interactions. If the effects of such motives is to minimize negative self-perceptions and maximize positive self-perceptions (particularly in non-deliberative circumstances), and if behaviors follow from such self-perceptions, then such processes may work to increase the emission of prosocial behaviors and decrease the emission of anti-social behaviors.

However, such effects are for future research to document. The take-home point from the present chapter is a simple one. In circumstance that do not promote deliberative processing of behavior, the actions of the self-enhancement/self-protection motives work to dampen the effects of a priming manipulation on interpretations of self-behavior and consequent self-judgments. However, when circumstances promote deliberative processing, a priming manipulation can directly affect interpretations of self-behavior and consequent self-judgments, bypassing the usual action of these motives.
References


