



## CHAPTER 22

# Classic Self-Deception Revisited

DELROY L. PAULHUS  
ERIN BUCKELS

In common parlance, *self-deception* is the act of lying to oneself. When more rigorous definitions are attempted, this straightforward notion quickly becomes complex, if not impossibly paradoxical. Especially problematic is the assumption that self-deception is analogous to deceiving others. These difficulties have undermined the feasibility of operationalizing the concept and conducting research. Despite this checkered history, the possibility of confirming the existence self-deception remains so seductive that we enter the fray one more time.

Rather than clarifying the concept, the theoretical literature on self-deception has become more abstruse and unwieldy. Recent treatises and responses thereto confirm the lack of consensus (e.g., McKay & Dennett, 2009; Mele, 1997). Instead, we see a diversity of competing but overlapping and intertwining concepts. Blatant challenges include recommendations to replace self-deception with concepts such as bad faith (Sartre, 1943/1982), disavowal of engagement (Fingarette, 1969), wishful thinking (Szabados, 1985), akrasia (Pears, 1984), positive illusions (Taylor & Brown, 1988), and the adaptive unconscious (Wilson, 2002).

The advent of psychology's cognitive revolution encouraged the study of more empirically tractable concepts such as confirmation bias (Kunda, 1990), self-signaling (Mijovic-Prelec & Prelec, 2010), dynamic complexity (Paulhus & Suedfeld, 1988), ethical fading (Tenbrunsel & Messick, 2004), cognitive avoidance (Greenwald, 1997), self-serving processing (von Hippel, Lakin, & Shakarchi, 2005), and time discounting (Ainslie, 1997). As he had done with other psychological conundra, Zajonc (1980) opened the door with evidence that affect and cognition follow independent routes to awareness. Thus, the information-processing approach easily incorporates the dual processing implicit in self-deception.

Serious consideration of this plethora of conceptual competitors is beyond the mandate of this chapter. Instead, a brief summary of the key issues will have to

suffice (the first section of the chapter). We then turn to several influential operationalizations that have led most directly to empirical investigation. Ultimately, we single out one of these—the Quattrone–Tversky paradigm—to generate and explain new data on the curious phenomenon of cheating on practice tests (the second section). Finally, we summarize and critique the individual-difference literature (the third section).

## Overview of the Issues

### *Everyday Self-Deception*

We begin by pointing to a common but puzzling observation: People sometimes appear to believe something that they must know is false. As observers, we conclude that such behavior is self-deceptive because the contradictory evidence should be obvious to all—especially the perpetrator. This phenomenon goes well beyond exaggeration, faking, or simple lying: In those cases, the individual is aware of uttering a falsehood. Instead, self-deception seems to be something deeper and more complicated—even paradoxical.

Consider some anecdotal examples. An otherwise pleasant young man is clearly an alcoholic but bristles at others saying so. He refuses to acknowledge the truth even though the evidence is obvious: There are empty bottles hidden throughout his apartment, and his boss has often sent him home for drinking on the job. Again, it doesn't count as self-deception if he knows he's an alcoholic but is simply lying about it.

Consider another case, where a young woman has a deep-seated hatred of her mother but cannot admit it to herself. The signs of this hatred are abundant: She angers quickly at any mention of her mother and makes a face when forced to discuss her. Perhaps the young woman cannot admit her feelings because much guilt and shame would surely ensue. What kind of horrible person hates her mother—the one who brought her into this world and sacrificed to raise her?

Consider the father who cannot believe that his only son, his pride and joy, is actually guilty of the heinous crimes alleged by the police. Because his reason for living would be shattered, the father cannot allow himself to believe the accusations. But he winces at every ring of the phone, fearing that it is the police calling (once again) about his son's latest misdeed.

The more carefully one analyzes such cases, the fuzzier the notion of self-deception becomes. The apparent duality of thought seems to implicate some process equivalent to the psychoanalytic unconscious: That traditional framework easily accommodates cases where an emotional conflict influences an individual's behavior while remaining inaccessible. At a conscious level, details about the conflict are unavailable or, at least, obscure. The unconscious, however, "knows" the truth. This terminology has proved so useful in framing complex behavior that it is now part of everyday lay conversations (Westen, 1998).

Unfortunately, such anecdotes—whether from personal experience or insights from clinicians—constitute the bulk of the evidence for the existence of self-deception. Before we describe the limited empirical literature, two other sources of evidence warrant mention.



### **Soft Versions of Self-Deception**

People are often inaccurate about aspects of their lives. But few of these self-inaccuracies implicate a self-deceptive process. Some follow from systematic limitations in our self-knowledge (Kihlstrom, 2001; Wilson, 1985). Others result from simple misinformation. For example, you may not have been told that you were adopted: In that case, you may have been purposely deceived by others. Or you may believe that you have a genius-level IQ because you accidentally miscored a take-home IQ test. Your recall of the fact that you hated your parents at age 10 may have faded along with other early memories. None of these cases qualifies as self-deception.

Other everyday phenomena seem to smack of self-deception but are less dramatic. They might be labeled as “soft” self-deception. You might set your watch 10 minutes ahead to ensure that you get to an appointment on time. How can that possibly work? You know very well your watch is 10 minutes fast. Yet people say it helps them to be on time—perhaps because the initial conclusion that they are late shocks them into action. Or take procrastination: Although the strategy has never paid off before, we again put off making that unpleasant phone call. Our intellectual powers allow us to devise impressive rationalizations for staying longer in bed or waiting until the last minute to write a paper.

It is a stretch to label such cases as self-deception. They are better placed into the category of crude coping mechanisms. When confronted, perpetrators immediately acknowledge the facts. The term *self-deception* should be reserved for cases where strong psychological forces prevent us from acknowledging a threatening truth about ourselves.

### **Evolutionary Basis**

But isn't truth distortion inherently maladaptive? Not so, according to a growing number of writers taking an evolutionary perspective. They argue that human beings engage in self-deception because it is built into the genes of our species (Lockard & Paulhus, 1988; Trivers, 1985). According to evolutionary theory, such psychological tendencies were gradually insinuated into our genetic makeup. The genes of individuals lacking this mechanism suffered a reproductive disadvantage (Krebs & Denton, 1997; Krebs, Denton, & Higgins, 1988).

But how could such irrationality be adaptive? One argument is that complete awareness of our motives would interfere with effectively satisfying them (Trivers, 1985). Believing that something is true facilitates its coming to fruition (Starek & Keating, 1991; von Hippel & Trivers, 2011). For example, confidence in making the Olympic team is a necessary (but not sufficient) condition for making the dream come true. Fending off an attacker is facilitated by exaggerated confidence. In both cases, however, there are negative consequences to inaccuracy: In the first case, one may delay reproduction with 4 years of futile workouts; in the second case, one may unnecessarily put life at risk. Overall, however, there is good reason to believe that a modicum of self-deception is adaptive (Baumeister, 1993).

Such arguments set the stage for debates over the existence of self-deception. If evolutionarily coherent, then self-deception does not have to be viewed as a human aberration (Lockard & Paulhus, 1988).

### ***Theoretical History and Background***

Self-deception is often discussed in the context of psychoanalytic theory. Rather than one of many defense mechanisms, self-deception is thought to be a necessary component of all psychoanalytic defenses (Sackeim, 1988). Each one entails the paradoxical element noted earlier: There must be at least one moment of self-deception for the successful operation of a defense mechanism. Readers familiar with defenses such as projection, intellectualization, and repression will understand the argument that, in each case, a person has to be both unaware and hyperaware of the disturbing information (Lockie, 2003; Westen, 1998).

Cases where the biased belief is maintained under continual confrontation might be labeled *deep self-deception*. As in the psychoanalytic notion of repression, some conflicts may quickly be resolved in favor of the psychologically comfortable option: Subsequently, an anxiety signal triggers an avoidance process that helps maintain the contradictory state with no further confrontation.

As a result, psychoanalytic theory has always been pessimistic about the possibility of people recognizing their own self-deceptions. However, clinicians and nonclinicians alike believe that insight is possible at some point down the line ("I must have been self-deceived about that relationship"). Such insight may not be possible until strong affect has subsided and a more objective analysis is possible. This psychoanalytic perspective continues to have broad appeal (see Paulhus, Fridhandler, & Hayes, 1997).

### ***Skepticism***

Over 70 years ago, Freud's ideas about self-deception were attacked by the celebrated philosopher, Jean-Paul Sartre. Along with many nonphilosophers, Sartre (1943/1982) dismissed the idea of self-deception as impossible. How can one know something and not know it at the same time?

This criticism continues to resonate with some contemporary commentators (Gergen, 1997; Kihlstrom, 2001; Szabados, 1985). They remain dubious about the very existence of deep self-deception. Some of this criticism smacks of residual antipathy toward psychoanalysis. Other writers seem concerned that the label of self-deception lets perpetrators off the moral hook (Fingarette, 1969). Other critics are simply skeptical about a processing system that requires monitoring and management of potentially upsetting self-knowledge. Continuous monitoring would require significant cognitive resources. Such cases, they argue, are better described as suppression, mental control, self-regulation, or repetitive coping (for a variety of sources, see Wegner & Pennebaker, 1993). Again, this suppression task would seem downright impossible without some awareness of the threatening thought.

### ***Information Processing***

Surprising to some early critics was the fact that the feasibility of self-deception has been supported by some of the most rigorous research in cognitive psychology. There is no longer any debate about the fact that many processes operate without awareness.



Moreover, we now know that the human cognitive apparatus allows for multiple representations of the same stimulus: Indeed, the modular nature of the brain allows for contradictory information to be stored in two different parts of the brain (Fodor, 1983). Finally, we also know that emotional aspects of a stimulus are processed more quickly than is the content. For example, it has been shown that, with a polygraph, the emotional impact of a word can be detected before the word is understood consciously (Zajonc, 1980).

Given the solid evidence for these cognitive processes, the possibility of self-deception becomes quite feasible. Several of these dual-process theories have been reviewed by Chaiken and Trope (1999; see also Gawronski & Bodenhausen, Chapter 3, this volume). In all these theories, one process deals with the informational content of the stimulus; the other with the emotional content. Moreover, the emotional system operates more quickly, thereby allowing the mind to set up preemptive roadblocks for the unacceptable. Greenwald's (1997) "junk-mail" analogy nicely explains how disturbing evidence can be overlooked indefinitely by responding with avoidance tactics to prevent any prospect of it surfacing.

### *Beyond Motivated Cognition*

With this body of research in mind, Mele (1997) attempted to "deflate" the language of self-deception by arguing that all such phenomena can be reduced to various types of motivated cognition (Kunda, 1990). Others (including yours truly) disagree. A satisfactory explanation of self-deception requires not only motivated cognition but also the additional feature of discrepant representations. Its uniqueness may follow from the direct involvement of the self. To aid in their decision, readers should explore the review of motivated cognition by Helzer and Dunning (Chapter 23, this volume).

## **Experimental Inductions**

So far we have established that (1) clinicians are confident that certain of their patients have deceived themselves, (2) self-deception is compatible with evolutionary psychology, and (3) the human information-processing system allows for self-deception.

On the empirical side, small advances have been made using a variety of operationalizations. For example, researchers have concluded that self-deception is easier in the future than in the present (Robinson & Ryff, 1999) and more likely in private than in public conditions (Smith & Whitehead, 1993). It appears only when a credible excuse is available (Ditto & Lopez, 1992) and when feedback about accuracy is vague (Slooman, Fernbach, & Haggmayer, 2010).

Nonetheless, direct experimental evidence for the existence of self-deception is hard to come by. Of course, it takes just one valid demonstration to confirm the possibility. But such demonstrations have proved to be challenging even in tightly controlled laboratory studies. In fact, only a handful of studies (described below) claim to have confirmed an instance of self-deception. We leave it to the reader to decide whether these studies are convincing or not.

### **Gur and Sackeim**

Harold Sackeim and Ruben Gur are credited with developing systematic criteria for self-deception (Sackeim & Gur, 1978), as well as laying out experimental evidence for its existence (Gur & Sackeim, 1979). Their criteria for demonstrating self-deception included (1) evidence for simultaneous but contradictory beliefs, only one of which is conscious, and (2) evidence that this discrepancy is motivated.

Their experimental paradigm built on the fact that people typically find the sound of their own voices to be aversive. Subjects were asked to distinguish their own from a series of voices and indicate so by saying "me" or "not me." Under certain conditions, subjects failed to acknowledge hearing their own voices.

At the same time subjects were hooked up to a polygraph that measured emotional responses to the voices. In short, the experimental setup provided two sources of information about whether the subjects recognized their own voices—an oral response, and an emotional response measured by polygraph.

Although the polygraph invariably showed a blip in reaction to the subject's own voice, the oral response was often inaccurate. In particular, *false denials were common*: Subjects said "not me" even though the polygraph invariably detected their voices. Such false denials suggested a bifurcation of awareness: The individual knows something and does not know it at the same time.

The researchers also predicted systematic changes in false denials depending on the need to avoid self-confrontation. This need was induced in some subjects by threatening their self-esteem using a standard failure induction. Indeed, after subjects' self-esteem was threatened, the frequency of false denials increased.

According to Sackeim and Gur (1978), false denials indicate that subjects believe *X* and do not believe *X* at the same time. Moreover, this lack of awareness is motivated by intrapsychic concerns (maintaining self-esteem). Together, the experiments appear to demonstrate an instance of self-deception. Of course, they do not address how widespread the phenomenon is. Instead, they point out that a single instance is all that is required to confirm that self-deception is possible.

The need for such elaborate laboratory orchestration is not without its critics. The artificiality and complexity of the Sackeim–Gur paradigm does little to convince many observers that self-deception is at large in everyday life.

### **Quattrone and Tversky**

The second study claiming to demonstrate self-deception was conducted by psychologists George Quattrone and Amos Tversky (1984). Their theoretical rationale was based on the common belief that changing one's score on any correlated variable can change a significant outcome. For example, assume for the moment that intellectuals tend to smoke pipes. The individual who believes that smoking a pipe will make him into an intellectual is being self-deceptive. Given that doctors are typically the source of bad news about one's health, some individuals choose to avoid doctors. They, too, are engaging in self-deception. Put another way, people often believe that correlation guarantees causation.

Although such beliefs are common, they are more likely (or more irrational) under motivated conditions. As such, Quattrone and Tversky attempted to demonstrate that



the experience of pain, which was self-evident under control conditions, could be minimized if its long-term implications were sufficiently serious.

### *The Pain Paradigm*

The experimenters exploited a cold pressor test, where participants are asked to immerse one hand in very cold water and keep it there “as long as you can stand it.” Before taking the test, some of the participants were threatened with information about an unfortunate correlate of the pain experience: “People who feel a lot of pain from the cold water have a weakness in their cardiovascular system. This defect leads to early heart attacks and a short lifespan.”

Results showed that participants receiving this information rated the task as less painful than did a control group. They even held their hand in the cold water longer. They seemed to be trying to convince themselves that they didn’t have the life-threatening cardiovascular defect. They were engaging in self-deception, according to Quattrone and Tversky, because they wouldn’t acknowledge—even to themselves—the pain that they surely were experiencing.

A potential concern that participants were simply engaging in impression management. Specifically, they may have minimized their pain reports in the health threat condition to avoid the public embarrassment. The researchers went to great lengths to minimize this possibility. For example, they used two experimenters: one to provide information about health consequences, and another to collect pain reports. See Quattrone and Tversky (1984) for other arguments to counter accusations of impression management.

Another concern was that actual pain, not just its perception, was being modified by such manipulations. This concern was overcome in a conceptual replication recently conducted by Sloman and colleagues (2010). Instead of modifying pain reports, the researchers showed that dot-tracking speed could be increased or decreased by telling participants that tracking speed was associated with higher or lower intelligence. The fact that control over tracking speed is immediate and conscious was used to argue that this motivated behavior was a form of self-deception.

### **Reflection on the Key Experiments**

At this point, readers may or may not be convinced that these studies demonstrate self-deception. What they should be convinced of is that confirmation of self-deception is incredibly difficult. Remember that a convincing experiment has to show that an individual believes something and disbelieves it at the same moment. The Sackeim-Gur (1978) study appears to have accomplished this goal directly; the Quattrone-Tversky (1984) study provides only indirect evidence of contradictory beliefs. On the other hand, the latter paradigm better captures everyday self-deception and permits any number of follow-up studies to explore the theoretical issues.

### **Cheating On Practice Tests**

The most recent research exploiting the Quattrone-Tversky paradigm comes from our own laboratory. We have conducted a series of studies on the phenomenon of

cheating on practice tests (Buckels & Paulhus, 2011; Paulhus, Nathanson, & Lau, 2005). Although common, this paradoxical behavior seems to meet the Quattrone–Tversky (1984) criteria for self-deception.

To prepare for an important test, it is common for students to take practice versions of the actual test. Preparation courses and handbooks for the SAT, GRE, MCAT, and so forth, do a thriving business and all include practice tests. On the surface, the purpose of practice tests is to diagnose one's current capability in this domain: They are designed to give test takers a realistic sense of where they stand and how much effort remains to reach their desired level of competence.

My discussions with Kaplan employees indicate that clients frequently cheat on the practice tests. In our pretests, we found that more than one-third of students take more than the allotted time on a practice test for the GRE. But if the purpose is realistic self-assessment, then why would anyone cheat on a practice test?

The answer appears to be self-deception for the sake of maintaining a positive self-evaluation. This positivity can be assured with a high score on the practice test. As Quattrone and Tversky (1984) argued, people believe that any change to a predictor will change the outcome: They have confused diagnostic contingencies with causal contingencies.<sup>1</sup> In this case, people seem to believe that even an underhanded way of improving practice test scores will improve their ability to perform better on ability tests in the future. By implication, such improvements will lead to greater scholastic success.

## **Our Studies**

Our laboratory has recently generated several studies designed to confirm this form of self-deception. Following the Quattrone–Tversky (1984) model, we stepped up the motivation for subjects to perform well on (what were described as) GRE practice tests. We also provided them with an opportunity to cheat by taking extra time. As far as participants could tell, such cheating would go undetected. At the same time, the ambiguity of taking a test on-line would permit subjects to believe the altered results.

### **Study 1**

Participants were randomly assigned to complete a vocabulary test under one of three conditions: (1) a self-enhancement condition, where the test was described as predicting future life success; (2) a reward condition, where a prize of \$200 was awarded for a top score; or (3) a control condition. They were advised that the time limit was 8 minutes: This countdown time appeared on the screen in full view at all times. However, the experimenter did not stop participants after 8 minutes.

Despite the fact that participants were well aware of the time limit and of the elapsed time, 37% of participants cheated by taking longer than the allotted time. As expected, participants in the self-enhancement condition displayed the greatest amount of cheating, that is, significantly longer test times than those in the reward and control conditions. Those who cheated were thereby unable to score higher on the test.



Directly after completing the task, participants were asked, "Why did you take extra time?" Those who had taken extra time justified their behavior with comments such as "I was compensating for being distracted" or "Otherwise my score would not represent my true ability." Overall, their message was that the cheated score was more accurate.<sup>2</sup> In short, the motivated subjects had come to believe in their obviously fudged scores.

### *Study 2*

To confirm that a behavior is motivated, the researcher must show that its direction can be reversed (Martin & Tesser, 2009; Sackeim & Gur, 1978). Otherwise, the motivation condition may have induced some other effect (e.g., arousal or distraction) that could explain the extra time taken. Hence, we needed to establish that the effect could be reversed. But is it possible to motivate student participants to perform worse than normal?<sup>3</sup>

To this end, we added a self-handicapping condition to Study 2. This induction was not easily effected: We had to convince educated subjects that doing well on an ability test was indicative of a negative outcome. Eventually, we found one that had some credibility. Drawing on the stereotype that genius and insanity often go hand in hand, we informed subjects about research linking high test scores to the style of creativity often found in people with schizophrenia.

Once again, participants in the self-enhancement condition cheated most (and achieved the highest scores). Participants in the self-handicapping condition cheated least (and scored lowest).

Apparently, self-handicappers found ways to sabotage their own performance! We investigated this apparent sabotage in several ways. First we examined the pattern of work times. Compared to the control condition, the self-handicapping group spent significantly less time working on the items. In fact, their median duration was actually less than the allotted 8 minutes. Not only their overall score but also the proportion of correct answers was lower. Perhaps their effort was reduced; perhaps they purposely put down wrong answers.

See the summary of conditions and results in Table 22.1. Together, these results indicate the self-deceptive nature of cheating on practice tests. Subjects can be motivated to excel or fail on a test that is supposed to inform them about their current ability level. The design and outcomes of our studies are consistent with the Quattrone–Tversky paradigm in indicating that such behavior is part of a more general (and apparently nonconscious) tendency purposely to bias predictors in the direction of desired outcomes.

### *Bottom Line*

Arguably, the key controversy over self-deception centers on whether individuals can ever hold two thoughts that are both simultaneous and contradictory. The consciously proclaimed belief (e.g., "My future is secure") should defer to overwhelming evidence that contradicts it. In the studies reviewed here, subjects' proclaimed beliefs were challenged by direct experience of their pain (Quattrone & Tversky, 1984), by

**TABLE 22.1. Summary of Inductions and Outcomes in Our Cheating Research**

Induction	Instructions	Test time taken	Test score outcome
<u>Study 1</u>			
1. Enhancement	Good performers go on to successful lives.	8.5 minutes	Highest
2. Monetary reward	Reward of \$200 for best performance.	8.3 minutes	Intermediate
3. Control group	No instructions.	8.1 minutes	Lowest
<u>Study 2</u>			
1. Enhancement	Good performers go on to successful lives.	8.5 minutes	Highest
2. Handicapping	People who do well are prone to schizophrenia.	7.9 minutes	Lowest
3. Control group	No instructions.	8.1 minutes	Intermediate

observing their contradictory behavior (Sloman et al., 2010), or by observing their own cheating (Paulhus et al., 2005). In all three cases, perpetrators “should” have been aware of the contradictory evidence. Ultimately, this criterion may rest on moral as well as factual grounds (Paulhus, Fridhandler, & Hayes, 1997).

### Individual Differences In Self-Deception

The research cited so far has focused on situational inducements/opportunities to engage in self-deception. As early as Frenkel-Brunswik (1939), other writers have focused on individual differences in self-deception. As in the conceptual debates, the focal measurement contrast is self-deception versus other-deception. In other words, the positive bias observed in self-descriptions may be aimed at the self or at others. Cronbach (1945), for example, made this distinction while discussing biases in questionnaire responses. The K scale developed for the Minnesota Multiphasic Personality Inventory (MMPI) was specifically targeted at self-deception. However, the first instrument specifically distinguishing self-deception and other-deception did not arrive until Sackeim and Gur (1978). That measure exhibited a coherent pattern of correlations with the MMPI validity scales.

Later, Paulhus (1984) refined the Gur–Sackeim scales while confirming their distinctiveness. To avoid the connotation of outright lying, the label *other-deception* was replaced with *impression management*. Later psychometric analyses of the Self-Deception Scale prompted the separation of two subscales: Self-Deceptive Enhancement (SDE) and Self-Deceptive Denial (SDD) (Paulhus & Reid, 1991). A further distinction was needed to acknowledge the difference between agentic and communal forms of self-deception (Paulhus, 2002; Paulhus & John, 1998).

The SDE scale taps a form of narcissistic exaggeration and outperforms impression management scales in predicting ego-relevant outcomes. Its correlates include tendencies toward overclaiming (Paulhus, Harms, Bruce, & Lysy, 2003), hindsight



bias (Paulhus, 1998b), and other self-favoring biases (Hoorens, 1995). High-SDE individuals also exhibit a discordance with reality, as indicated by a discrepancy in self-ratings of agency relative to ratings by group consensus (Paulhus, 1998a). Nonetheless, SDE correlations with concrete performance are invariably positive (Johnson, 1995; Starek & Keating, 1991). More recently, SDE has also shown utility in moderating the validity of other self-report scales (Otter & Egan, 2007). For example, controlling for SDE served to improve the validity (defined by the correlation with observer ratings) of self-report neuroticism (Berry, Page, & Sackett, 2007).

Indirect operationalizations of individual differences have also been used. One example is the tendency to engage in self-serving bias (SSB) (Ditto & Lopez, 1992; von Hippel et al., 2005). Ditto and Lopez established its cross-situational consistency and showed that the SSB could predict a behavioral index of cheating<sup>4</sup> (von Hippel et al., 2005). Not surprisingly, such indexes are less commonly used than questionnaire measures.

How many types of self-deceptive tendencies are there? Paulhus and John (1998) addressed this question by factoring a set of 60 self-criterion discrepancies calculated from a broad range of personality and ability variables. They isolated two large factors, namely, agentic and communal enhancement. This two-factor model has been replicated by others (Honkaniemi & Feldt, 2008; Lönnqvist, Verkasalo, & Bezmenova, 2007; Vecchione & Alessandri, in press).

### Limitations

From the beginning, difficulties in interpretation have dogged social desirability scales (Paulhus, 2002). For the most part, however, the critiques have been directed at impression management scales, based primarily on their overlap with actual positive traits (e.g., Uziel, 2010). To date, we are unaware of any critiques leveled at individual-difference measures of self-deception.

Unfortunately, only modest headway has been achieved in linking the individual-difference research with the definitional criteria for self-deception addressed in the prominent experimental paradigms. The task of diagnosing contradictory beliefs within the same questionnaire has proved especially daunting. Instead, self-deceptive tendencies have been investigated as predictable aspects of established traits (e.g., narcissism) and cultural styles (Lalwani, Shavitt, & Johnson, 2006).

We can point to several research programs that do capture the spirit of dual representation. Instead of the label *self-deception*, those programs have applied terms such as *repression* and *defensiveness* (see Paulhus et al., 1997). The work on repressive style exemplifies the contrast between behavior collected at different levels of awareness (e.g., Bonanno & Singer, 1990; Coifman, Bonanno, Ray, & Gross, 2007; Weinberger, 1990). The extensive body of work on the dynamics of narcissism includes demonstrations of distinctive reactions of narcissists under threat and no-threat conditions (for a review, see Morf & Rhodewalt, 2001). Finally, the dualism model is evident in work contrasting implicit and explicit self-esteem (Jordan, Spencer, Zanna, Hoshino-Browne, & Correl, 2003): Of the four combinations, most fragile is the individual with high explicit and low implicit self-esteem.

Such individuals tend to overreact to criticism by exhibiting extreme negative affect (Kernis & Goldman, 2006). All of these empirically based approaches have supported the conclusion that some individuals more than others manage to maintain multiple representations of the same information—and that one of those representations induces sufficient distress to minimize its full availability to conscious awareness.

## Conclusions

Clinicians are certain about the operation of self-deception in some of their patients. Given its complexity, it is not surprising that only a handful of empirical studies claim to have confirmed the phenomenon. Instead, the bulk of writing on self-deception has been published by philosophers who—unlike psychologists—do not have to collect data to support their claims. Based on the work reviewed here, our position is that self-deception research has advanced sufficiently to confirm its existence as a genuine human phenomenon.

Most convincing to us is the research based on the Quattrone–Tversky paradigm, which provides a simple but powerful framework for pursuing the otherwise elusive phenomenon of self-deception. Several laboratories, including our own, have exploited this paradigm and provided conceptual replications of the original operationalization. Using a variety of labels, the search for an effective individual-difference measure continues among personality researchers.

## Implications

The notion of self-deception implicates a deep-seated psychological process that eventuates in a distorted self-perception. As in the case of other motivated biases, the victim possesses the information to draw the correct conclusion but, for emotional reasons, does not do so. Self-deception goes further to suggest that both the accurate and inaccurate representations remain active. To regulate dangerous information most effectively, one must, at some level, recognize and manage it. As such, it may be seen as the most extreme version of motivated bias.

The phenomenon of self-deception may play an important, if hidden, role in a variety of human endeavors. Despite its evolutionary roots, the long-term impact of self-deception in the modern world appears to be predominantly toxic (see Leary & Toner, Chapter 25, this volume). It places limitations on the ability of humans to stave off social and political conflict. It impairs any rational approach to financial prosperity. Its psychological mechanism entails an intrapsychic pressure to trade off or, at least, balance two fundamental motivations: People seek accurate information about their world and its complexity; they also need to defend against information that would undermine the values and ideas on which their identities are constructed. It remains to be seen whether our vital intrapsychic lies are best interpreted in moral or factual terms.



## NOTES

1. Of course, changing some predictors does benefit future performance: for example, studying the dictionary, reviewing high-school mathematics. These behaviors actually do play a causal role in improving one's abilities, as well as future test scores.
2. Note the contrast with purposeful watch setting: When people are reminded that they purposely set their watch ahead, they immediately acknowledge the real time.
3. Such self-handicapping was successfully induced in recent studies by Sloman et al. (2010).
4. Unfortunately, their self-deception index did not correlate significantly with SDE.

## REFERENCES

- Ainslie, G. (1997). If belief is a behavior, what controls it? *Behavioral and Brain Sciences*, 20, 102–104.
- Baumeister, R. F. (1993). Lying to yourself: The enigma of self-deception. In M. Lewis & C. Saarni (Eds.), *Lying and deception in everyday life* (pp. 166–183). New York: Guilford Press.
- Berry, C. M., Page, R. C., & Sackett, P. R. (2007). Effects of self-deceptive enhancement on personality–job performance relationships. *International Journal of Selection and Assessment*, 15, 94–109.
- Bonanno, G. A., & Singer, J. L. (1990). Repressive personality style: Theoretical and methodological implications for health and pathology. In J. L. Singer (Ed.), *Repression and dissociation: Implications for personality theory, psychopathology, and health* (pp. 435–470). Chicago: University of Chicago Press.
- Buckels, E., & Paulhus, D. L. (2011, June). *Cheating on practice tests: A matter of self-deception?* Paper presented at the meeting of the Canadian Psychological Association, Toronto.
- Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford Press.
- Coifman, K. G., Bonanno, G. A., Ray, R. D., & Gross, J. J. (2007). Does repressive coping promote resilience?: Affective–autonomic response discrepancy during bereavement. *Journal of Personality and Social Psychology*, 92, 745–758.
- Cronbach, L. J. (1945). Response sets and test validity. *Educational Psychological Measurement*, 6, 475–494.
- Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and non-preferred outcomes. *Journal of Personality and Social Psychology*, 63, 568–584.
- Fingarette, H. (1969). *Self-deception*. New York: Humanities Press.
- Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Frenkel-Brunswik, E. (1939). Mechanisms of self-deception. *Journal of Social Psychology*, 10, 409–420.
- Gergen, K. J. (1997). Detecting self-deception. *Behavioral and Brain Sciences*, 20, 114–115.
- Greenwald, A. G. (1997). Self-knowledge and self-deception: Further consideration. In M.S. Myslobodsky (Ed.), *The mythomanias: The nature of deception and self-deception* (pp. 51–71). Mahwah, NJ: Erlbaum.
- Gur, R. C., & Sackeim, H. A. (1979). Self-deception: A concept in search of a phenomenon. *Journal of Personality and Social Psychology*, 37, 147–169.

- Honkaniemi, L., & Feldt, T. (2008). Egoistic and moralistic bias in real-life inventory responses. *Personality and Individual Differences*, *45*, 397–311.
- Hoorens, V. (1995). Self-favoring biases, self-presentation, and the self–other asymmetry in social comparison. *Journal of Personality*, *63*, 793–817.
- Johnson, E. A. (1995). Self-deceptive responses to threat: Adaptive only in ambiguous circumstances. *Journal of Personality*, *63*, 759–791.
- Jordan, C. H., Spencer, S. J., Zanna, M. P., Hoshino-Browne, E., & Correl, J. (2003). Secure and defensive high self-esteem. *Journal of Personality and Social Psychology*, *85*, 969–978.
- Kernis, M. H., & Goldman, B. M. (2006). A multicomponent conceptualization of authenticity: Theory and research. *Advances in Experimental Social Psychology*, *38*, 283–357.
- Kihlstrom, J. F. (2001). Dissociative disorders. In P. B. Sutker & H. E. Adams (Eds.), *Comprehensive handbook of psychopathology* (3rd ed., pp. 259–276). New York: Plenum Press.
- Krebs, D. L., & Denton, K. (1997). Social illusions and self-deception: The evolution of biases in person perception. In J. A. Simpson & D. T. Kenrick (Eds.), *Evolutionary social psychology* (pp. 21–47). Hillsdale, NJ: Erlbaum.
- Krebs, D. L., Denton, K., & Higgins, N. (1988). On the evolution of self-knowledge and self-deception. In K. McDonald (Ed.), *Sociobiological perspectives on human behavior* (pp. 103–139). New York: Springer-Verlag.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.
- Lalwani, A. K., Shavitt, S., & Johnson, T. (2006). What is the relation between cultural orientation and socially desirable responding? *Journal of Personality and Social Psychology*, *90*, 165–178.
- Lockie, R. (2003). Depth psychology and self-deception. *Philosophical Psychology*, *16*, 127–148.
- Lockard, J. S., & Paulhus, D. L. (Eds.). (1988). *Self-deception: An adaptive mechanism?* Englewood Cliffs, NJ: Prentice-Hall.
- Lönnqvist, J. E., Verkasalo, M., & Bezmenova, I. (2007). Agentic and communal bias in socially desirable responding. *European Journal of Personality*, *21*, 853–868.
- Martin, L. L., & Tesser, A. (2009). Five markers of motivated behavior. In G. B. Moskowitz & H. Grant (Eds.), *The psychology of goals* (pp. 257–276). New York: Guilford Press.
- McKay, R. T., & Dennett, D. C. (2009). The evolution of misbelief. *Behavioral and Brain Sciences*, *32*, 493–561.
- Mijovic-Prelec, D., & Prelec, D. (2010). Self-deception as self-signaling: A model and experimental evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*, 227–240.
- Mele, A. R. (1997). Real self-deception. *Behavioral and Brain Sciences*, *20*, 91–136.
- Morf, C. C., & Rhodewalt, F. (2001). Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model. *Psychological Inquiry*, *12*, 177–196.
- Otter, Z., & Egan, V. (2007). The evolutionary role of self-deceptive enhancement as a protective factor against antisocial cognitions. *Personality and Individual Differences*, *43*, 2258–2269.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, *46*, 598–609.
- Paulhus, D. L. (1986). Self-deception and impression management in test responses. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaire* (pp. 143–165). New York: Springer-Verlag.
- Paulhus, D. L. (1998a). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology*, *74*, 1197–1208.



- Paulhus, D. L. (1998b). *Manual for Balanced Inventory of Desirable Responding (BIDR-7)*. Toronto: Multi-Health Systems.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 67–88). Hillsdale, NJ: Erlbaum.
- Paulhus, D. L., Harms, P. D., Bruce, M. N., & Lysy, D. C. (2003). The over-claiming technique: Measuring self-enhancement independent of ability. *Journal of Personality and Social Psychology*, *84*, 681–693.
- Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic bias in self-perceptions: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, *66*, 1024–1060.
- Paulhus, D. L., Nathanson, C., & Lau, K. S.-L. (2005, August). *Cheating on practice tests: A matter of self-deception?* Paper presented at the annual meeting of the American Psychological Society, New York.
- Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in social desirable responding. *Journal of Personality and Social Psychology*, *60*, 307–317.
- Paulhus, D. L., & Suedfeld, P. (1988). A dynamic complexity model of self-deception. In J. S. Lockard & D. L. Paulhus (Eds.), *Self-deception: An adaptive mechanism?* (pp. 132–145). Englewood Cliffs, NJ: Prentice-Hall.
- Paulhus, D. L., Fridhandler, B., & Hayes, S. (1997). Psychological defense: Contemporary theory and research. In J. A. Johnson, R. Hogan, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 543–579). New York: Academic Press.
- Pears, D. (1984). *Motivated irrationality*. Oxford, UK: Clarendon Press.
- Quattrone, G. A., & Tversky, A. (1984). Causal versus diagnostic contingencies: On self-deception and the voter's illusion. *Journal of Personality and Social Psychology*, *46*, 236–248.
- Robinson, M. D., & Ryff, C. D. (1999). The role of self-deception in perceptions of past, present, and future happiness. *Personality and Social Psychology Bulletin*, *25*, 595–606.
- Sackeim, H. A. (1988). Self deception: A synthesis. In J. S. Lockard & D. L. Paulhus (Eds.), *Self deception: An adaptive mechanism?* (pp. 146–165). Englewood Cliffs, NJ: Prentice-Hall.
- Sackeim, H. A., & Gur, R. C. (1978). Self-deception, other-deception and consciousness. In G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research* (Vol. 2, pp. 139–197). New York: Plenum Press.
- Sartre, J.-P. (1982). *Mauvaise foi* and the unconscious. In R. Wollheim & J. Hopkins (Eds.), *Philosophical essays on Freud*. Cambridge, MA: Cambridge University Press. (Original work published 1943)
- Slooman, S. A., Fernbach, P. M., & Hagmayer, Y. (2010). Self-deception requires vagueness. *Cognition*, *115*, 268–281.
- Smith, S. H., & Whitehead, G. I. (1993). The use of self-deceptive strategies as a function of the manipulation of publicness: The role of audience information. *Contemporary Social Psychology*, *17*, 8–13.
- Starek, J. E., & Keating, C. F. (1991). Self-deception and its relationship to success in competition. *Basic and Applied Social Psychology*, *12*, 145–152.
- Szabados, B. (1985). The self, its passions, and self-deception. In M. W. Martin (Ed.), *Self-deception and self-understanding* (pp. 143–168). Lawrence: University Press of Kansas.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.
- Tenbrunsel, A. E., & Messick, D. M. (2004). Ethical fading: The role of self-deception in unethical behavior. *Social Justice Research*, *17*, 223–236.

- Trivers, R. (1985). *Social evolution*. Menlo Park, CA: Benjamin/Cummings.
- Uziel, L. (2010). Rethinking social desirability scales: From impression management to interpersonally oriented self-control. *Perspectives on Psychological Science*, 5, 243–262.
- Vecchione, M., & Alessandri, G. (in press). Alpha and beta traits and egoistic and moralistic self-enhancement: A point of convergence between two research traditions. *Journal of Personality*.
- von Hippel, W., Lakin, J. L., & Shakarchi, R. J. (2005). Individual differences in motivated social cognition: The case of self-serving information processing. *Personality and Social Psychology Bulletin*, 31, 1347–1357.
- von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences*, 34, 1–56.
- Wegner, D. M., & Pennebaker, J. W. (Eds.). (1993). *Handbook of mental control*. Englewood Cliffs, NJ: Prentice-Hall.
- Weinberger, D. A. (1990). The construct validity of the repressive coping style. In J. L. Singer (Ed.), *Repression and dissociation: Implications for personality theory, psychopathology, and health* (pp. 337–386). Chicago: University of Chicago Press.
- Westen, D. (1998). The scientific legacy of Sigmund Freud: Towards a psychodynamically informed psychological science. *Psychological Bulletin*, 124, 333–371.
- Wilson, T. D. (1985). Self-deception without repression: Limits on access to mental states: In M. W. Martin (Ed.), *Self-deception and self-understanding* (pp. 95–116). Lawrence: University Press of Kansas.
- Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Harvard University Press.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151–175.