

## Accepted Manuscript

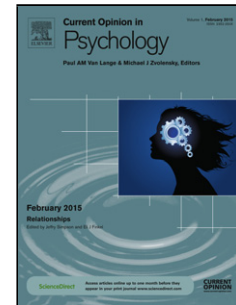
Title: The What and Why of Self-Deception

Author: Zoë Chance Michael I. Norton

PII: S2352-250X(15)00186-4

DOI: <http://dx.doi.org/doi:10.1016/j.copsyc.2015.07.008>

Reference: COPSYP 137



To appear in:

Received date: 19-3-2015

Accepted date: 26-7-2015

Please cite this article as: Z. Chance, M.I. Norton, The What and Why of Self-Deception, *COPSYP* (2015), <http://dx.doi.org/10.1016/j.copsyc.2015.07.008>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## The What and Why of Self-Deception

### Highlights

- We compare three types of definitions of self-deception
- We present three hypotheses regarding the adaptive nature of self-deception
- We review recent empirical research providing evidence for all three

Running Head: SELF-DECEPTION

The What and Why of Self-Deception

Zoë Chance

Yale University

Michael I. Norton

Harvard University

Correspondence should be addressed to Zoë Chance, Yale School of Management, 165 Whitney Ave., New Haven, CT 06511, [zoe.chance@yale.edu](mailto:zoe.chance@yale.edu), phone (857) 891-5013, fax (203) 815-1557.

### Abstract

Scholars from many disciplines have investigated self-deception, but both defining self-deception and establishing its possible benefits have been a matter of heated debate – a debate impoverished by a relative lack of empirical research. Drawing on recent research, we first classify three distinct definitions of self-deception, ranging from a view that self-deception is synonymous with positive illusions to a more stringent view that self-deception requires the presence of simultaneous conflicting beliefs. We then review recent research on the possible benefits of self-deception, identifying three adaptive functions: deceiving others, social status, and psychological benefits. We suggest potential directions for future research.

Word count: 1,604

The nature and definition of self-deception remains open to debate. Philosophers have questioned whether – and how – self-deception is possible; evolutionary theorists have conjectured that self-deception may – or must – be adaptive [1-3]. Until recently, there was little evidence for either the existence or processes of self-deception; indeed, Robert Trivers [4] (p. 15) wrote that research on self-deception is still in its infancy. In recent years, however, empirical research on self-deception has been gaining traction in social psychology and economics, providing much-needed evidence and shedding light on the psychology of self-deception [5]. We first classify competing definitions of self-deception, then review recent research supporting three distinct advantages of self-deception: improved success in deceiving others, social status, and psychological benefits.

### **What is self-deception?**

While definitions of self-deception abound, most can be grouped into one of three broad categories (acknowledging that these generalizations ignore many nuanced distinctions). In the first definition – sometimes called “deflationary” [2] – self-deception is simply a motivated false belief [e.g. 5,6], and is indistinguishable from positive illusions. In this view, self-deception can arise from, for example, from selective attention, biased information search, or forgetting. In the second definition, self-deception is a motivated false belief that persists in spite of disconfirming evidence [e.g. 7,8]. In this view, not all positive illusions are self-deceptive, and biased information search does not qualify as self-deception since disconfirming evidence is not encountered and ignored; however, self-deception can still arise via selective attention or

knowledge avoidance [9,10], where disconfirming evidence is present – but ignored or discounted.

In the final definition [e.g. 11,12,13], self-deception is a motivated and conscious false belief held simultaneously with a conflicting unconscious true belief. The paradox inherent in this view of self-deception – how can a person believe  $p$  and simultaneously convince himself *not*  $p$ , while still retaining the original belief? – has led some scholars to declare self-deception to be impossible [2]. While many studies have investigated the first two definitions of self-deception, there are critical barriers to assessing this final definition: it requires both observing the unconscious true belief and confirming that the stated or expressed belief is sincere. In the classic demonstration of this class of self-deception [11], participants' stated denials of recognizing their own recorded voices were belied by (unconscious) skin conductance measures.

Note that in all cases, however, self-deception is presumed to arise from a motivated desire to see the self and the world in ways that favor the self. As with the many definitions of self-deception, scholars have also examined the benefits of self-deception using a wide array of perspectives. We next review recent research that suggests self-deception may be adaptive in (at least) three distinct ways: to deceive others, to gain social rewards, and to reap psychological benefits.

### **Deceiving the self to deceive others**

Self-deception may have evolved as an adaptive strategy for deceiving others without being discovered [4,5]. In this context, self-deception can prevent the liar from emitting nonverbal cues of guilt, minimize the cognitive load associated with lying, and reduce retribution via pleas of ignorance. Skeptics of this theory, however, have noted that not only would a

disregard for the truth solve most of these problems without the very real costs of inaccuracy [14,15], but that most examples of self-deception do not involve deceiving other people [16].

Recent studies provide compelling evidence for a relationship between self-deception and other-deception [17]. Given that high-status individuals are both more likely to be the targets of deception and better equipped to punish deceivers they catch in the act, due to their control of resources, it was hypothesized that self-deception may be more likely to occur when lying to high-status than equal-status individuals. Participants first studied a set of words; next, when these words were presented along with new words, they declared verbally that they had not seen any of the words while privately pressing a key to indicate whether they had seen each word or not; importantly, participants were face-to-face with either a high-status (“teacher”) or equal-status (“student”) confederate, and were told they would lose their payment if the confederate detected their dishonesty. When the confederate left the room, participants again privately indicated which words had been part of the original group. Participants demonstrated poorer private recall when they lied to the teacher than to the student—evidence that they had temporarily forgotten the truth and self-deceived.

### **Social benefits of self-deceptive confidence**

Confidence impacts the way others perceive and treat us. Confident people are believed more often [e.g. 18], have more influence in groups or when giving advice [19,20], and appear more knowledgeable [21]. Most germane to our review, inaccurate (and self-deceptive) overconfidence often carries social benefits: in one study, overconfidence increased peers’ judgments of competence, resulting in higher social status [25]. Similarly, overconfidence correlates with others’ mistakenly high expectations: following an unstructured interaction in

preparation for a class, students predicted their next grades and relative ranking in the group and the grades and relative rankings of others. Self-deceptive overconfidence was positively correlated with fooling others: students who had overly high expectations for themselves inspired overly high expectations from others [22]. Note, however, that as being perceived as lying to others is costly, so being perceived as self-enhancing can also come with costs: whereas actual self-enhancers were rated by others as emotionally stable, socially attractive, and socially influential, those who observers also *perceived* to be high self-enhancers (those whom observers “saw through”) were perceived as emotionally unstable and socially unstable – though they remained socially influential [23]. These results align with the “optimal margin of illusion” perspective [24]. Self-deception can come with social benefits, but those benefits are not guaranteed.

### **Psychological benefits of self-deception**

Finally, recent research suggests that people may derive direct utility from self-deceptive beliefs—viewing themselves, others, the state of the world, or their future prospects in preferred and self-serving ways [25-27]. As just one example, not only do optimists work harder to achieve their expected results, they gain direct utility from optimism [28].

Moreover, self-deception induced in the moment can also have psychological benefits. In one set of studies [29], some participants had an opportunity to cheat on a test by looking at an answer key. They then systematically overpredicted their performance on future tests: rather than attributing their prior achievement to the presence of the answers, they convinced themselves they would do well even without the answer key. When paid for both accuracy and performance, they earned less money as a result of their self-deceptive optimism when they could not meet



their own high expectations. In another investigation, participants who self-deceived about their performance on a judgment task in order to bolster false hope of winning bonus money were also quicker to make those self-deceptive judgments than other judgments—suggesting a hasty cover-up of their suspicion about the truth [30].

Relatedly, self-deceptive individuals understate the degree to which motivation affects their behavior, to reinforce their motivated beliefs that their actions are diagnostic of a desired but unobservable trait. These studies build on a classic experiment [31] in which individuals held their hands in cold water longer when endurance was purported to indicate good health – participants persisted longer on a task (enduring a painful finger prick or searching for hidden pictures) when told persistence augured well (good skin in the future or self-control in other domains such as eating) [32]. Furthermore, persistence corresponded with lower reported effort. This effort denial allowed participants to benefit from the belief that their persistence was due to their underlying desirable type rather than to their high motivation.

### **Conclusion & future directions**

We have offered a classification of current views of self-deception and summarized research demonstrating three distinct benefits of deceiving the self. Still, despite the recent increase in research exploring the topic, more research is needed to understand two critical aspects: when self-deception occurs, and how to stop it.

First, it is not yet known whether self-deception occurs during the encoding or the retrieval process – although there is some suggestive evidence for both [33]. Consider the “cheating on tests” paradigm [31], in which participants who performed well on a test by glancing at the answers at the bottom believed that their performance was due to their ability

rather than the answers. Did this self-deception happen as they viewed the answers, whereby as soon as they glanced at an answer they simultaneously attributed getting the question right to their ability? Or does the awareness of having used the answers linger – for seconds, or even milliseconds – and only then is wiped away? Functional MRI could be used build on new work [34], further illuminating the processes of self- versus other-deception.

Second, research documenting interventions to help prevent self-deception are clearly needed. Ambiguity has been shown to be necessary for self-deception [35], but there are likely other factors in addition to clarity of evidence that may prevent self-deception. One promising avenue is to explore the decay rate of self-deception when exposed to the truth. In one investigation [Chance, Gino, Norton & Ariely, unpublished], people who had cheated on a test (and deceived themselves into thinking they did well due to their ability) took a series of additional tests without access to the answers; self-deception lingered beyond the next test, but, faced with repeated poor performance, people eventually stopped believing they were better than they actually were.

The “what” of self-deception – what distinguishes self-deception from other phenomena? – and the “why” of self-deception – what adaptive functions can it serve? – have been of interest to scholars from a variety of disciplines. The recent proliferation of research has begun to shed new light, but social scientists must be careful not to deceive ourselves into thinking that the psychology is fully understood.

## References

- [1]Pears DF: **Motivated irrationality**. 1984.
- [2]Mele AR: **Understanding and explaining real self-deception**. Behavioral and Brain Sciences 1997, **20**:127-134.
- [3]Bandura A: **Self-deception: A paradox revisited**. Behavioral and Brain Sciences 2011, **34**:16-17.
- [4]Trivers R: **The folly of fools: The logic of deceit and self-deception in human life**: Basic Books; 2011.
- [5]Von Hippel W, Trivers R: **The evolution and psychology of self-deception**. Behavioral and Brain Sciences 2011, **34**:1-16.
- [6]Mele A: **Self-deception and delusions**. European Journal of Analytic Philosophy 2006, **2**:109-124.
- [7]Bok S: **Secrets: On the ethics of concealment and revelation**: Vintage; 1989.
- [8]Mitchell J: **Living a lie: Self-deception, habit, and social roles**. Human Studies 2000, **23**:145-156.
- [9]Greenwald AG: **Self-knowledge and self-deception: Further consideration**. The mythomanias: The nature of deception and self-deception 1997:51-72.
- [10]Sharot T, Korn CW, Dolan RJ: **How unrealistic optimism is maintained in the face of reality**. Nature neuroscience 2011, **14**:1475-1479.
- [11]Gur RC, Sackeim HA: **Self-deception: A concept in search of a phenomenon**. Journal of Personality and Social Psychology 1979, **37**:147.
- [12]Audi R: **Self-deception, action, and will**. Erkenntnis 1982, **18**:133-158.
- [13]Pinker S: **Representations and decision rules in the theory of self-deception**. Behavioral and Brain Sciences 2011, **34**:35-37.
- [14]McKay R, Mijović-Prelec D, Prelec D: **Protesting too much: Self-deception and self-signaling**. Behavioral and Brain Sciences 2011, **34**:34-35.
- [15]Dunning D: **Get thee to a laboratory**. Behavioral and Brain Sciences 2011, **34**:18-19.
- [16]Khalil EL: **The weightless hat: Is self-deception optimal?** Behavioral and Brain Sciences 2011, **34**:30-31.
- [17]Lu HJ, Chang L: **Deceiving yourself to better deceive high-status compared to equal-status others**. Evolutionary Psychology 2014, **12**:635-654.
- [18]Penrod S, Cutler B: **Witness confidence and witness accuracy: Assessing their forensic relation**. Psychology, Public Policy, and Law 1995, **1**:817.
- [19]Swol LM, Snizek JA: **Factors affecting the acceptance of expert advice**. British Journal of Social Psychology 2005, **44**:443-461.
- [20]Zarnoth P, Snizek JA: **The social influence of confidence in group decision making**. Journal of Experimental Social Psychology 1997, **33**:345-366.
- [21]Price PC, Stone ER: **Intuitive evaluation of likelihood judgment producers: Evidence for a confidence heuristic**. Journal of Behavioral Decision Making 2004, **17**:39-57.
- [22]Lamba S, Nityananda V: **Self-Deceived Individuals Are Better at Deceiving Others**. PloS one 2014, **9**:e104562.
- [23]Dufner M, Denissen J, Sedikides C, Van Zalk M, Meeus WH, Aken M: **Are Actual and Perceived Intellectual Self-enhancers Evaluated Differently by Social Perceivers?** European Journal of Personality 2013, **27**:621-633.
- [24]Baumeister RF: **The optimal margin of illusion**. Journal of Social and Clinical Psychology 1989, **8**:176-189.
- [25]Kunda Z: **The case for motivated reasoning**. Psychological bulletin 1990, **108**:480.

- [26]Hsee CK: **Elastic justification: How tempting but task-irrelevant factors influence decisions.** Organizational Behavior and Human Decision Processes 1995, **62**:330-337.
- [27]Norton MI, Vandello JA, Darley JM: **Casualty and social category bias.** Journal of personality and social psychology 2004, **87**:817.
- [28]Foster G, Frijters P: **The formation of expectations: Competing theories and new evidence.** Journal of Behavioral and Experimental Economics 2014, **53**:66-81.
- [29]Chance Z, Norton MI, Gino F, Ariely D: **Temporal view of the costs and benefits of self-deception.** Proceedings of the National Academy of Sciences 2011, **108**:15655-15659.
- [30]Mijović-Prelec D, Prelec D: **Self-deception as self-signalling: a model and experimental evidence.** Philosophical Transactions of the Royal Society B: Biological Sciences 2010, **365**:227-240.
- [31]Quattrone GA, Tversky A: **Causal versus diagnostic contingencies: On self-deception and on the voter's illusion.** Journal of Personality and Social Psychology 1984, **46**:237.
- [32]Fernbach PM, Haggmayer Y, Sloman SA: **Effort denial in self-deception.** Organizational Behavior and Human Decision Processes 2014, **123**:1-8.
- [33]Anderson MC, Hanslmayr S: **Neural mechanisms of motivated forgetting.** Trends in cognitive sciences 2014, **18**:279-292.
- [34]Farrow TF, Burgess J, Wilkinson ID, Hunter MD: **Neural correlates of self-deception and impression-management.** Neuropsychologia 2015, **67**:159-174.
- [35]Sloman SA, Fernbach PM, Haggmayer Y: **Self-deception requires vagueness.** Cognition 2010, **115**:268-281.

## Annotated references

## References

- [1]Pears DF: **Motivated irrationality**. 1984.
- [2]Mele AR: **Understanding and explaining real self-deception**. Behavioral and Brain Sciences 1997, **20**:127-134.
- [3]Bandura A: **Self-deception: A paradox revisited**. Behavioral and Brain Sciences 2011, **34**:16-17.
- [4]Trivers R: **The folly of fools: The logic of deceit and self-deception in human life**: Basic Books; 2011.
- [5]Von Hippel W, Trivers R: **The evolution and psychology of self-deception**. Behavioral and Brain Sciences 2011, **34**:1-16.
- [6]Mele A: **Self-deception and delusions**. European Journal of Analytic Philosophy 2006, **2**:109-124.
- [7]Bok S: **Secrets: On the ethics of concealment and revelation**: Vintage; 1989.
- [8]Mitchell J: **Living a lie: Self-deception, habit, and social roles**. Human Studies 2000, **23**:145-156.
- [9]Greenwald AG: **Self-knowledge and self-deception: Further consideration**. The mythomanias: The nature of deception and self-deception 1997:51-72.
- [10]Sharot T, Korn CW, Dolan RJ: **How unrealistic optimism is maintained in the face of reality**. Nature neuroscience 2011, **14**:1475-1479.
- [11]Gur RC, Sackeim HA: **Self-deception: A concept in search of a phenomenon**. Journal of Personality and Social Psychology 1979, **37**:147.
- [12]Audi R: **Self-deception, action, and will**. Erkenntnis 1982, **18**:133-158.
- [13]Pinker S: **Representations and decision rules in the theory of self-deception**. Behavioral and Brain Sciences 2011, **34**:35-37.
- [14]McKay R, Mijović-Prelec D, Prelec D: **Protesting too much: Self-deception and self-signaling**. Behavioral and Brain Sciences 2011, **34**:34-35.
- [15]Dunning D: **Get thee to a laboratory**. Behavioral and Brain Sciences 2011, **34**:18-19.
- [16]Khalil EL: **The weightless hat: Is self-deception optimal?** Behavioral and Brain Sciences 2011, **34**:30-31.
- \*[17]Lu HJ, Chang L: **Deceiving yourself to better deceive high-status compared to equal-status others**. Evolutionary Psychology 2014, **12**:635-654.

This is the first paper we are aware of providing convincing evidence for a relationship between self-deception and lying. Participants lie to themselves (not recognizing stimuli they have seen before) when lying to high-status others. When not lying to others, they again remember having seen these stimuli.

- [18]Penrod S, Cutler B: **Witness confidence and witness accuracy: Assessing their forensic relation**. Psychology, Public Policy, and Law 1995, **1**:817.
- [19]Swol LM, Snizek JA: **Factors affecting the acceptance of expert advice**. British Journal of Social Psychology 2005, **44**:443-461.
- [20]Zarnoth P, Snizek JA: **The social influence of confidence in group decision making**. Journal of Experimental Social Psychology 1997, **33**:345-366.
- [21]Price PC, Stone ER: **Intuitive evaluation of likelihood judgment producers: Evidence for a confidence heuristic**. Journal of Behavioral Decision Making 2004, **17**:39-57.

- \*[22]Lamba S, Nityananda V: **Self-Deceived Individuals Are Better at Deceiving Others**. PloS one 2014, **9**:e104562.

These authors find that other people have high expectations of those who (privately) have high expectations of themselves. Whether these individuals are self-deceived is debatable, and there is no reason to believe they were intentionally misleading anyone else.

- \*[23]Dufner M, Denissen J, Sedikides C, Van Zalk M, Meeus WH, Aken M: **Are Actual and Perceived Intellectual Self-enhancers Evaluated Differently by Social Perceivers?** European Journal of Personality 2013, **27**:621-633.

Actual self-enhancers are deemed emotionally stable, socially attractive, and socially influential. However, being perceived as self-enhancing bears costs: perceived high self-enhancers were perceived as emotionally unstable and socially unstable – though they remained socially influential.

- [24]Baumeister RF: **The optimal margin of illusion**. Journal of Social and Clinical Psychology 1989, **8**:176-189.
- [25]Kunda Z: **The case for motivated reasoning**. Psychological bulletin 1990, **108**:480.
- [26]Hsee CK: **Elastic justification: How tempting but task-irrelevant factors influence decisions**. Organizational Behavior and Human Decision Processes 1995, **62**:330-337.
- [27]Norton MI, Vandello JA, Darley JM: **Casusitry and social category bias**. Journal of personality and social psychology 2004, **87**:817.
- \*[28]Foster G, Frijters P: **The formation of expectations: Competing theories and new evidence**. Journal of Behavioral and Experimental Economics 2014, **53**:66-81.

This paper discusses an econometric analysis comparing grade expectations to outcomes in a large dataset from multiple Australian universities. Controlling for demographic, psychographic, and many other background factors, the authors find a bias toward overly optimistic expectations, in patterns consistent with both anticipatory savoring and utility from high self esteem--which may provide further utility by allowing confident people to be convincing to others regarding their own merits.

- [29]Chance Z, Norton MI, Gino F, Ariely D: **Temporal view of the costs and benefits of self-deception**. Proceedings of the National Academy of Sciences 2011, **108**:15655-15659.
- [30]Mijović-Prelec D, Prelec D: **Self-deception as self-signalling: a model and experimental evidence**. Philosophical Transactions of the Royal Society B: Biological Sciences 2010, **365**:227-240.
- [31]Quattrone GA, Tversky A: **Causal versus diagnostic contingencies: On self-deception and on the voter's illusion**. Journal of Personality and Social Psychology 1984, **46**:237.
- \*[32]Fernbach PM, Hagmayer Y, Sloman SA: **Effort denial in self-deception**. Organizational Behavior and Human Decision Processes 2014, **123**:1-8.

These studies build on Quattrone & Tversky's (1984) classic experiment in which individuals held their hands in cold water longer when endurance was purported to indicate good health – participants persisted longer on a task (enduring a painful finger prick or searching for hidden pictures) when told persistence augured well (good skin in the future or self-control in other domains such as eating). Furthermore, persistence corresponded with lower reported effort.

This effort denial allowed participants to benefit from the belief that their persistence was due to their underlying desirable type rather than to their high motivation.

[33]Anderson MC, Hanslmayr S: **Neural mechanisms of motivated forgetting**. Trends in cognitive sciences 2014, **18**:279-292.

\*[34]Farrow TF, Burgess J, Wilkinson ID, Hunter MD: **Neural correlates of self-deception and impression-management**. Neuropsychologia 2015, **67**:159-174.

In an fMRI study, the authors tested neural activation during faking "good" or "bad" responses to the Impression Management and Self-Deceptive Enhancement sections of the Balanced Inventory of Desirable Responding (Paulhus, 1994), compared to a letter-counting baseline. The authors find that responses on the self-deceptive enhancement scale are less intentionally malleable, and that this process is neurally distinguishable from impression management--however, it is questionable whether this study captures self-deception.

[35]Sloman SA, Fernbach PM, Hagmayer Y: **Self-deception requires vagueness**. Cognition 2010, **115**:268-281.