CONCEPT

Peer-reviewed | Open Access

The Psychology of Self-Deceit: Why We Lie to Ourselves

Sarah Chi*

St. Clement's School, 21 St Clements Ave, Toronto, ON M4R 1G8, Canada

ARTICLE INFO

*Corresponding Author
Sarah Chi
E-mail: chi.sarahc@gmail.com

ABSTRACT

Self-deceit is a psychological paradox in which individuals convince themselves of false beliefs despite contradictory evidence. While long debated, the mechanisms and purposes of self-deceit remain unsettled. Methods: This conceptual paper synthesizes experimental and theoretical literature, drawing on intentionalist and deflationary models, alongside empirical studies, to evaluate mechanisms and functions of self-deceit. Results: Findings suggest self-deceit operates as an unconscious process influenced by biases, emotions, and desires. A combined model better explains the paradox, supported by evidence from psychophysiological and cognitive load experiments. Three primary purposes are identified: enhancing deception toward others, preserving self-esteem, and serving as a defensive mechanism for self-preservation. Discussion: By merging intentionalist and deflationary perspectives, this analysis highlights the adaptive role of self-deceit in reducing cognitive load, maintaining psychological well-being, and facilitating social functioning. Conclusions: Self-deceit is best understood as a non-intentional psychological process in which conflicting beliefs are motivated by unconscious desires, simultaneously supporting self-validation, interpersonal deception, and ego protection. Understanding its mechanisms and purposes provides insights into human cognition, emotion, and behaviour.

Keywords: Cognitive Load, Defence Mechanisms, Intentionalist Approach, Self-Deceit, Self-Esteem, Unconscious Processes.

1. INTRODUCTION

Self-deceit occurs when an individual convinces themselves of a false belief despite evidence supporting the truth. An experiment by Gur and Sackeim (1979) attempted to uncover a more nuanced understanding of this concept. When asked to identify whether a recording was their own voice or that of a stranger, participants incorrectly stated that a recording was not their voice despite having a higher psychophysiological reaction that indicated recognition. These findings suggest that participants simultaneously held two contradictory beliefs: they unconsciously recognized their voice, yet claimed the opposite. This paradox has sparked a long-standing debate regarding self-deceit, and its purpose is continually being researched and understood (Chance & Norton, 2015). Consequently, studying self-deceit is crucial, as it may reveal how humans navigate internal conflicts and decisions. By examining its functions and processes, psychologists can better understand its connection to human emotions and behaviours. Numerous perspectives can be used to understand why we lie to ourselves and whether there is a benefit to this seemingly counterintuitive thinking (Warren, 2014).

This study aims to answer the question: What is self-deceit? Beyond that, it will provide a nuanced understanding of two topics: the intent and mechanisms behind self-deceit and the purpose of its existence. Both points aid in providing nuanced answers to this question. This essay describes why self-deceit is an unconscious process motivated by internal states and discusses the possible benefits of lying to oneself.

2. THEORETICAL FRAMEWORK AND METHOD

This article adopts a conceptual synthesis approach rather than an empirical design. The objective is to critically examine the phenomenon of self-deceit by integrating insights from philosophy and psychology to clarify both its mechanisms and its purposes.

To achieve this, the analysis draws upon two primary theoretical traditions: the intentionalist framework, which emphasizes the paradox of consciously deceiving oneself, and the deflationary framework, which treats self-deceit as a motivated but non-intentional cognitive process. These perspectives were selected because they represent the dominant approaches to understanding self-deceit and frame the long-standing debate within both philosophical and psychological literature.

The sources reviewed include seminal philosophical arguments, peer-reviewed psychological studies, and empirical experiments that illuminate cognitive and affective mechanisms (e.g., memory disruption, psychophysiological measures, and cognitive load). Selection of materials was guided by their scholarly influence, citation within the field, and relevance to the questions of "how" and "why" self-deceit occurs.

The analytic strategy involved thematic categorization of the literature into two dimensions: (1) the mechanisms underlying self-deceit, and (2) the purposes it serves for individuals and social interactions. By comparing and contrasting intentionalist and deflationary perspectives with supporting psychological evidence, the article develops a synthesized model that highlights the adaptive and protective roles of self-deceit.

This methodological orientation allows for a nuanced conceptual contribution: rather than presenting new data, it refines theoretical understanding, clarifies debates, and points toward directions for future empirical research.

3. FINDINGS

3.1 Findings on Mechanisms of Self-Deceit

Deceit generally involves one party that is aware of the truth and a second party that has been provided with false information (Makowski et al., 2024). In contrast, self-deceit implies that an individual is both the deceiver and the deceived, creating a paradox. The definition of self-deceit as a "false belief held simultaneously with a contrasting unconscious true belief" (Chance & Norton, 2015) is highly debated by psychologists and philosophers. However, this contradiction can be explained by defining self-deceit as a conflict between unconscious and conscious beliefs motivated by biases and emotional reasoning. Two significant psychological models must be considered to understand this contradiction: the

intentionalist and deflationary(non-intentionalist) approaches (Chance & Norton, 2015). While both approaches attempt to address the paradox of conflicting perceptions, it is the combination of the two theories that defines self-deceit.

The intentionalist approach asserts that self-deceit is purposeful and acknowledges its paradoxical nature. In order to account for this contradiction, some theorists believe the mind is partitioned, meaning that two or more sections of the mind can hold two conflicting notions simultaneously (Mele, 2018). This explains how an individual can have opposing beliefs (Mele, 2018). This theory is further supported by a study which observed that self-deceit was more cognitively taxing, hypothesized to be a result of the state of conflict within one's mind when deceiving oneself (Li & Liu, 2024); however, this theory has been critiqued because it suggests that one section of the mind deceives another, while most minds are not fragmented (Mele, 2018). Alternatively, American philosopher Robert Audi (1982) claims that self-deceit is a state in which someone holds an opposing unconscious belief that is being deliberately ignored in favour of a false one. Psychologists Newman and Erber (2002) provide the example of a t Soldier who recognizes the immorality of his actions; to deceive himself into believing he is not as involved in these actions; he must hide the truth and evidence supporting unconsciously to maintain this misbelief. However, this model still leads to paradox; specifically, Mele (2018) notes that it does not address what he calls the dynamic paradox: If a person knows the truth, how can they intentionally deceive themselves into believing the opposite? With the example of the Nazi soldier, it seems implausible that he could intentionally deceive himself into believing he was not involved in the actions of Nazi Germany while simultaneously knowing the truth of his clear involvement in such activities. When an individual is aware of their deception, it becomes ineffective; therefore, it is irrational for one person to effectively convince themselves of a false notion if they are both the deceiver and the deceived (Winchell, 2015).

On the other end of this spectrum is the deflationary approach, which suggests that self-deceit is more similar to falsely or mistakenly believing something rather than holding two contradictory beliefs (Mele, 2001). By simplifying the nature of self-deceit, non-intentionalists avoid the paradox altogether. Instead, Mele (2001) argues that this inaccurate perception is neither intentional nor accidental but motivated by desire. Mele's definition is comparable to wishful thinking or motivated reasoning, as it lacks the conflict and tension that exists in the intentionalist viewpoint (Chance & Norton, 2015). However, philosophers such as Audi (1982) and Bermúdez (2000) critique Mele's perspective, arguing that a crucial

aspect of deception is the recognition of evidence that opposes their beliefs. For example, a person who denies that their partner is cheating despite evidence supporting the contrary often feels high stress and tension because of this conflicting evidence. Even when they maintain the belief that their partner remains loyal, suspicions may still manifest, even if subconsciously. Overall, while the deflationary approach accounts for the numerous paradoxes of self-deception, it lacks the tension characteristic of those experiencing self-deceit. These limitations suggest that a more comprehensive framework is needed to address these paradoxes.

Neither the intentionalist nor the deflationary approach alone is sufficient to explain the complexity of self-deceit. This concept is better understood as a combination of the two models. This definition of self-deceit involves contradictory beliefs; however, the process is not intentional but is driven by desires, emotions, and unconscious biases. By maintaining that a person can hold two opposing viewpoints, one conscious and one unconscious, this definition accounts for the conflict present in self-deceit and highlights that an individual is at least somewhat aware of evidence opposing their conscious beliefs. Moreover, empirical evidence further suggests that self-deceit arises unconsciously rather than through intent. Through an experiment comparing the cognitive loads of self-deceptive and non-deceptive behaviour, psychologists found that the cognitive load for the self-deceptive group was lower (Jian et al., 2019). Further experiments revealed that this reduction in cognitive load was connected to Involuntary Conscious Memory (ICM), which involves the recollection of "unintentional and spontaneous" (Jian et al., 2019, para. 6) memories or experiences. Essentially, ICM refers to the evocation of memories in the mind without conscious effort. Studies have highlighted that the ICM involuntarily disrupts the memories of participants (Jian et al., 2019). These findings highlight that self-deceit is not a result of intention but of memory disruption in the ICM to reduce cognitive load. Self-deceit is non-intentional, but rather one's desires and emotions act similarly in motivating the act of deception.

Because there is no intent to deceive, this definition effectively addresses the dynamic paradox. Ultimately, this definition of self-deceit most effectively explains the mechanisms underlying this concept.

3.2 Findings on Purposes of Self-Deceit

Deception is commonly observed in humans for social and relational purposes, such as avoiding conflict with others and increasing trust through false statements (Makowski et al., 2024). Additionally, it is crucial to understand the purpose of self-deceit to better understand the human mind. However, while numerous experiments study the process of self-deceit, the question of why it occurs remains unanswered. Although lying to others is seen as a daily occurrence that plays a crucial role in social settings (Li & Liu, 2024), lying to oneself may seem counterintuitive, as it creates false narratives that can harms individuals.

However, by examining multiple psychological perspectives, self-deceit can be defined in terms of three functions: deceiving others, self-validation, and serving as a protective mechanism.

A widespread explanation for self-deceit is that it helps deceive others more effectively (Chance & Norton, 2015). To ensure that one is not giving away verbal cues or nervous signals, a liar deceives themselves into believing the lie that they are telling (von Hippel & Trivers, 2011). Recent studies have emphasized the relationship between self-deceit and other-deceit. In an experiment, students were made to either lie to their teacher (who had a higher status than the participant) or lie to a fellow student (who had an equal status) (Jian et al., 2019). Self-deceit was observed more frequently in participants who lied to someone of higher status (Jian et al., 2019), which can be explained by the fact that lying to someone of high status often leads to more severe repercussions. The participants knew that if they were caught, they would be reprimanded, so by deceiving themselves, they could better deceive their teachers without being exposed. Alternate theories regarding self-deceit do not involve deceiving others but rather focus on how it validates oneself. Yang and colleagues (2024) describe self-deception as a response to a "comparative gap," finding it more common during upward social comparisons—such as when participants believed they performed worse in a game than others. These individuals also experienced stronger negative emotions and lower self-esteem, suggesting that self-deceit serves a self-preservative function (Yang et al., 2024).

A commonly proposed function of self-deceit is its role in preserving one's self-esteem. For example, participants who are provided with an answer sheet during tests may deceive themselves into believing that their performance on the test is a result of high intelligence, rather than recognizing the answer key as an advantage (Chance et al., 2011). By overestimating their own intelligence, self-deceit supports a person's positive view of themselves. Therefore, self-deceit validates an inflated sense of intelligence, even if it is only a short-term benefit (Chance et al., 2011). This experiment highlights the social and cognitive aspects of self-deceit and how deceiving oneself can benefit self-esteem. Additionally, self-deceit could benefit an individual's self-confidence in

public settings. Speaker anxiety has become a significant concern among researchers and professionals, which leads to a negative impact on the communication process (Bynum, 2025). To combat speaker anxiety, Bynum (2025) emphasizes prioritizing audience engagement over one's own internal negative thoughts. This effect could be described by self-deceit's role in the preservation of self-esteem: it allows speakers to better mask their anxiety and convince themselves of the audience's receptiveness to their words. In this sense, by maintaining one's self-esteem, self-deceit helps individuals speak more effectively and improves their self-assurance.

Finally, beyond inner validation, self-deceit can act as a defence mechanism in humans. Psychologist Goldeman (1985) examined a psychoanalytic perspective and how selectively ignoring certain facts and memories can act as a means of self-preservation. Sigmund Freud's ego defence mechanisms discuss the concept of denial, which is the refusal to believe that something is true (Warren, 2014). For example, people who claim to have a healthy lifestyle despite smoking regularly practice denial to feel less guilt about their habits. Another common mechanism is rationalization, which is used to justify one's thoughts and beliefs to maintain a positive view of oneself (Warren, 2014). These strategies are designed to protect the ego, the core rational sense of self (Warren, 2014). The function of self-deceit is like that of a defence mechanism. By keeping certain thoughts out of our conscious awareness, this process reduces stress and anxiety, protects self-esteem, and maintains social harmony (Goleman, 1985). While Freud emphasized defence mechanisms like denial and rationalization, today's individuals turn to digital escapism. Escapism, defined as the avoidance of the real world in favour of a virtual world, shields individuals from stressful and difficult moments in life, to the point that some "forget the physical world (Subudhi et al., 2020, p. 38)." The positive influence of self-deceit is short-term; these mechansims of defense can regulate an individual's mental health, subjective well-being, and interpersonal relationship (Liu et al., 2025). Overall, the psychoanalytic perspective highlights the short-term benefits of self-deceit in protecting one's sense of self.

Together, these explanations suggest that self-deceit is not an illogical flaw in human cognition but a form of adaptation, preservation, and protection. First, self-deceit evolved to more convincingly deceive others. The fact that self-deception is more likely when lying to those of a higher status emphasizes this s of self-deceit. Second, a more common purpose of self-deceit in our everyday lives is to preserve a positive view of oneself. The final explanation of the role of self-deceit answers why humans may repress certain memories or favour a belief over another; self-deceit is a protective mechanism that preserves the human ego. Therefore, self-deceit exists to benefit three purposes: lying to others, maintaining self-esteem, and protecting humans from difficult facts and memories.

4. DISCUSSION

This conceptual analysis demonstrates that self-deceit cannot be adequately explained by either intentionalist or deflationary accounts in isolation. Instead, the evidence supports a hybrid perspective, one that acknowledges unconscious bias, affective drivers, and selective memory as central to the phenomenon. Such a synthesis helps reconcile the paradox that individuals both recognize and deny contradictory evidence without the need for fully partitioned minds or wholly accidental misbeliefs.

From a broader psychological standpoint, the findings align with research on cognitive dissonance and motivated reasoning, which similarly illustrate how humans maintain coherence between beliefs and desires, even in the face of counter-evidence. Self-deceit can therefore be interpreted as one adaptive strategy within a larger repertoire of defensive cognitive processes, alongside rationalization and denial. Importantly, the reviewed evidence suggests that self-deceit reduces cognitive load, highlighting its potential role in conserving mental resources when individuals face threatening or disconfirming information.

The analysis also suggests that the purposes of self-deceit—to enhance deception of others, to validate the self, and to protect against psychological harm—reflect both social and intrapsychic functions. Socially, self-deceit increases the plausibility of lies by minimizing behavioural cues of insincerity. Intrapersonally, it preserves self-esteem and shields individuals from destabilizing realities, thereby promoting short-term resilience. However, these functions may come at a long-term cost if self-deceit prevents individuals from addressing maladaptive behaviours, interpersonal conflicts, or ethical lapses.

Conceptually, the integration of intentionalist and deflationary models provides a more complete explanation of self-deceit, but it also raises new questions. How can unconscious motivational processes be empirically distinguished from deliberate acts of denial? To what extent does self-deceit overlap with culturally shaped narratives of self-presentation or identity management? Future empirical studies might address these gaps by combining neurocognitive measures of memory suppression with behavioural paradigms of impression management, offering a more rigorous test of the boundaries of self-deceit.

Overall, this discussion situates self-deceit within the larger literature on human cognition and behaviour.

Rather than a paradoxical flaw in reasoning, self-deceit emerges as a flexible adaptation with both protective and deceptive functions. By reframing it in this way, the article underscores its relevance not only to philosophy and psychology but also to ethics, education, and clinical practice, where understanding the dynamics of self-deception may inform interventions that promote self-awareness and healthier decision-making.

4.1 Limitations and Future Directions

As a conceptual synthesis, this article does not present original empirical data. Its analysis is therefore limited to the scope of existing philosophical debates and psychological studies available in the literature. While this approach provides theoretical clarity, it cannot determine the prevalence, intensity, or variability of self-deceit across cultures, age groups, or social contexts. Another limitation is the reliance on classic psychological experiments which, though foundational, may not capture the full complexity of human cognition in contemporary settings. For example, while cognitive load studies observe that self-deceit reduces mental effort, these experiments are set in controlled environments that do not reflect natural experiences. Additionally, due to its theoretical nature, Freud's theory of defence mechanisms is limited in terms of its benefit to contemporary research.

Future research could address these gaps in several ways. First, **empirical studies** that integrate neurocognitive measures (e.g., brain imaging of memory suppression) with **behavioural paradigms** (e.g., impression management or lying tasks) would help distinguish unconscious motivational processes from deliberate acts of denial. Second, **cross-cultural investigations** could explore how cultural norms and moral frameworks shape the forms and functions of self-deceit. Finally, **applied studies** might examine the implications of self-deceit in domains such as education, mental health, and ethics, particularly in interventions designed to foster self-awareness, resilience, and healthier decision-making.

By acknowledging these limitations and outlining directions for future work, this article positions self-deceit not only as a philosophical puzzle but also as a rich avenue for interdisciplinary exploration in psychology, neuroscience, and the humanities.

5. CONCLUSION

In conclusion, this analysis sought to define self-deceit by addressing a paradox, presenting its mechanisms, and explaining the purpose of this concept. Self-deceit can be defined as simultaneously having an unconscious and conscious belief that is driven by an individual's affective drivers and internal biases; experimental evidence emphasizes the simultaneous occurrence of conscious and unconscious beliefs (Gur and Sackeim, 1979). While this process occurs unintentionally, factors like human emotion and desire often act in an intentional way, leading to the development of certain beliefs over others. These motivational forces also result in the selective ignorance of facts or memories.

Characterizing self-deceit as contradicting beliefs motivated by emotions and desire, this paper argues for a synthesis of the popular intentionalist and Deflationary approaches. This integrated perspective of both models offers a novel argument of self-deceit that advances beyond existing debates, addressing key conflicts and paradoxes in current literature. As a result, this definition sets the stage for a deeper examination of the purposes and implications of self-deceit.

This paper analyzed three different purposes of self-deceit to gain a full understanding of its existence in humans. The most common theory as to why self-deceit evolved in humans is that it aids in better deceiving others. Other instances of self-deception that do not involve deceiving others present the idea of maintaining a positive view of oneself by overlooking certain facts and memories. Through a psychodynamic lens, it can be seen as a defence mechanism, protecting an individual's mind by reducing their stress and anxiety, as self-deceit can repress harmful memories. Understanding self-deceit not only benefits psychologists but also has implications for daily behaviour and decision-making, providing insight into how humans navigate conflicting beliefs and respond to unwanted truths. To effectively understand the human mind, it is crucial to recognize how self-deception shapes human thoughts, beliefs, and choices.

Acknowledgements

Funding

Declaration of Conflicting Interests

The author declares that she has no competing interests.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

During the preparation of this manuscript, the author(s) did not employ any of the Generative AI and/or AI-Assisted technologies for Language refinement, drafting background section and did not perform any Task of the technology.

References

Audi, R. (1982). Self-deception, action, and will. *Erkenntnis (1975)*, *18*(2), 133–158. https://doi.org/10.2307/20010803

Bermúdez, J. L. (2000). Self deception, intentions, and contradictory beliefs. *Analysis*, *60*(4), 309–319. https://doi.org/10.2307/3329185

Bynum, D. (2025). *Targeting challenges to manage the fear of public speaking*. **Horizon Journal of Humanities & Social Sciences Research, 7**(1), 120–130. https://doi.org/10.37534/bp.jhssr.2025.v7.n1.id1284.p120.

Chance, Z., & Norton, M. I. (2015). The What and Why of self-deception. *Current Opinion in Psychology*, *6*(6), 104–107. https://doi.org/10.1016/j.copsyc.2015.07.008

Chance, Z., Norton, M. I., Gino, F., & Ariely, D. (2011). Temporal view of the costs and benefits of self-deception. *Proceedings of the National Academy of Sciences*, *108*(Supplement_3), 15655–15659. https://doi.org/10.1073/pnas.1010658108

Goleman, D. (1985). *Vital lies, simple truths: The psychology of self-deception*. Simon and Schuster Paperbacks. https://books.google.ca/books?hl=en&lr=&id=R_jF5yUjYSEC&oi=fnd&pg=PA11&ots=636vjAr__l3g&sig=ngXNkVNuISCRXBQI-0O1n7IgjUA&redir_esc=y#v=onepage&q&f=false

Gur, R. C., & Sackeim, H. A. (1979). Self-deception: A concept in search of a phenomenon. *Journal of Personality and Social Psychology*, *37*(2), 147–169. https://doi.org/10.1037/0022-3514.37.2.147

Jian, Z., Zhang, W., Tian, L., Fan, W., & Zhong, Y. (2019). Self-deception reduces cognitive load: the role of involuntary conscious memory impairment. *Frontiers in Psychology*, *10*(1718). https://doi.org/10.3389/fpsyg.2019.01718

Li, Y., Liu, Z. (2024). The effects of deception on memory: a comparative study of actors and eyewitnesses accounts. BMC Psychology, 12, 638. https://doi.org/10.1186/s40359-024-02152-3

Liu, J., Ding, W., Deng, L., Tan, M., & Guan, P. (2025). *The influence of different negative feedback on the decay of self-deception. Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1499089

Makowski, D., Lau, Z. J., Pham, T., Te, A. S., Kirk, S., Liauw, C., & Chen, S. H. A. (2024). *The heart can lie: a preliminary investigation of the role of interoception and theory of mind in deception. Current Psychology*, *43*(44), 34215–34226. https://doi.org/10.1007/s12144-024-06890-w

Mele, A. R. (2001). *Self-deception unmasked*. Princeton University Press. https://doi.org/10.2307/j.ctt7s4tg

Mele, A. R. (2018). Self-deception. *Routledge Encyclopedia of Philosophy, Taylor and Francis*. 10 https://doi.org/10.4324/9780415249126-v033-1

Newman, L. S., & Erber, R. (2002). *Understanding genocide: the social psychology of the Holocaust*. Oxford University Press.

Subudhi, R. N., Das, S. C., & Sahu, S. (2020). *Digital escapism*. **Horizon Journal of Humanities & Social Sciences Research, 2**(S), 37–44. https://doi.org/10.37534/bp.jhssr.2020.v2.nS.id1071.p37

von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences*, *34*(1), 1–16. https://doi.org/10.1017/s0140525x10001354

Warren, C. (2014). *Honest Liars: Using Psychological Theory to Understand Self-Deception*. Www.psichi.org. https://www.psichi.org/page/183EyeSprSum14aWarr

Winchell, D. S. (2015). Paradigmatic self-deception. *Arts & Sciences Electronic Theses and Dissertations*, *459*. https://doi.org/10.7936/k7r78cbj

Yang, Y., Zhong, B., Zhang, W., & Fan, W. (2024). The Effect of Comparative Direction and Comparative Gap on Self-Deception. *Psychology Research and Behavior Management*, *Volume 17*, 2819–2834. https://doi.org/10.2147/prbm.s467437

Biographical Statement of Author(s)

Sarah Chi is a high school student living in Toronto, Canada. She attends St. Clement's School. Her areas of interest include psychology, philosophy, and education.

Ms. Sarah Chi
St. Clement's School
21 St Clements Ave
Toronto, ON M4R 1G8
CANADA.
E-mail: chi.sarahc@gmail.com