Information Research - Vol. 30 No. CoLIS (2025)

Informational self-deception: deconstruction and reconstruction within philosophy of information

Juliana Mestre

DOI: *https://doi.org/10.47989/ir30CoLIS52312*

Abstract

Introduction. Hannah Arendt argues that, in the twentieth century, there arose diffuse systems of mendacity tied to information communication. Such '*modern lies*' engendered self-deception among both officials and citizens, allowing the perpetuation of atrocities. In this paper, I examine self-deception from within philosophy of information, confronting and deconstructing a central dichotomy in self-deception research: that between ontological and psychological renderings of the concept.

Method. Derridean deconstruction is used to create and mine tension between ontological and psychological self-deception. Ultimately, informational self-deception is reconstructed within this tension.

Analysis. Ontological self-deception positions deceit as a natural and passive condition of being-in-the-world. As members of an entropic infosphere, we are all subject to self-deception, and we thus have an imperative to continually question ourselves and our information. Psychological self-deception positions such deceit as an active concept; rather than the passive condition of the many, self-deception is the active choice of a few. This opens questions surrounding intentionality, motivation and variances in levels of deceit.

Conclusion. Ultimately, I reconstruct informational self-deception as sustained by the tensions between ontological and psychological self-deception, raising new questions for philosophy of information regarding what it means to be deceived.

Introduction: Hannah Arendt and the modern lie

There are, on a basic level, three entities involved in an act of deception. First, there is the deceiver, the one intentionally crafting a fabrication. Second, there is the lie itself, the content that is fabricated. Third, there is the person or group of people being deceived. In philosophy of information, significant attention has been granted to the first two. Don Fallis (2010; 2022), for example, questions what it means to deceive in the twenty-first century, challenging and expanding traditional definitions of the lie to account for more nuanced forms of deception. Additionally, theorists focus on fabricated content, establishing and challenging definitions for mis- and disinformation (Fallis, 2011, 2015, 2016), examining the impact of such mis- and disinformation in various political contexts (Palomo, 2021; Søe, 2018, 2021), and confronting traditional conceptions of information quality from different levels of abstraction (Floridi and Illari, 2014). Less attention, however, is paid to the third entity involved in an act of deception: the deceived. What does it mean to succumb to a specific mediatic and informational narrative despite evidence to the contrary? What are the political and ethical stakes of doing so?

These questions become increasingly complicated in more diffuse mediatic systems where a specific deceiver is difficult, perhaps impossible, to locate. When narratives are propagated by institutions, governments and corporations and spread from multiple outlets on multiple platforms both by individuals and by bots, then there exists a breakdown in the traditional formulation of deceiver and deceived. Hannah Arendt was a twentieth-century theorist interested in this breakdown between the traditional formulation of deceiver and deceived, and she dedicated considerable attention to unravelling what she deemed to be a transformation in what it means to lie. In my introduction, therefore, I turn to Arendt (1968) to establish what she calls '*the modern lie*' and to introduce the dangerous self-deception such mendacity engenders (p. 253).

Arendt was a German Jewish philosopher who fled Germany as Naziism gripped the country prior to WWII and who, throughout her career, dedicated considerable attention to unravelling diffuse systems of mendacity against the backdrop of the twentieth century. Specifically, she examined political events like the rise of totalitarian regimes (Arendt, 1973), the post-war trials of high-ranking Nazi officials (Arendt, 1963; 1971; 2006), and the production and reception of the Pentagon Papers in the US (Arendt, 1972). Her political and philosophical writings unearthed what she argued to a be a transformation of the lie from the traditional lie, specific and targeted towards one's enemy, to the modern lie, dispersed and often targeted towards one's own citizenry. On the one hand, '*the traditional lie concerned only particulars and was never meant to deceive literally everybody*', while, on the other hand, the modern lie '*is the relatively recent phenomenon of mass manipulation of fact and opinion as has become evident in the rewriting of history, in image making, and in actual government policy*' (Arendt, 1968, pp. 252-253). This modern lie is not spoken of in explicitly informational terms by Arendt, who died in 1975 at the very onset of an informational turn in theory. However, it is nonetheless an informational phenomenon, as such deception encompasses political systems of informational propaganda proliferated by increasingly widespread and sophisticated information and communication technologies (ICTs). The modern lie, diffuse and dangerous, is informational. Moreover, such mendacity is not a relic of the twentieth century; structures of information communication are more pervasive and advanced than ever, and they foster similar modern lies today (La Caze, 2017).

Though Arendt dedicates volumes to unpacking different case studies, characteristics and consequences of the modern lie in the twentieth century, there is one aspect of such deception that I find particularly intriguing for philosophy of information – the propensity of the modern lie to breed self-deception. She defines self-deception as,

> *the nature of our ability to deny in thought and word whatever happens to be the case. This active, aggressive capability is clearly different from our passive susceptibility to falling prey*

> *to error, illusion, the distortions of memory, and to whatever else can be blamed on the failings of our sensual and mental apparatus.* (Arendt, 1972, pp. 4-5).

While self-deception is not a new concept, Arendt (1968) argues that the modern lie is more likely than the traditional lie to produce self-deception even amongst those responsible for deceiving: '*the business of [political] deception […] used to belong to [those…] who among themselves still knew and could preserve the truth. […T]hey could deceive others without deceiving themselves*' (p. 253). In the twentieth century, however, '*self-deception [became] an indispensable tool in the trade of image-making*' and in the functioning of increasingly large bureaucratic and governmental apparatuses (Arendt, 1968, p. 253). It is to self-deception that Arendt (1963) attributes some of the most horrific atrocities of WWII, arguing that holocaust organisers like Adolf Eichmann had submitted to a banal evil by ceasing to question themselves or their actions, instead willingly falling prey to a thoughtless self-lie that their bureaucratic duties were necessary and justified rather than abhorrent and divorced from factuality. Moreover, she suggests that such self-deception was widespread amongst the German citizenry during the Third Reich: '*Eighty million Germans had been shielded against reality and factuality by exactly the same self-deception, lies, and stupidity that had now become ingrained in Eichmann's nature*' (Arendt, 1963). The ethical and political stakes of self-deception are thus set, and Arendt's texts pose a warning to twenty-first century scholars to be wary of self-deception, evolving forms of mendacity and the relationship between mendacity and power.

In this paper, I am interested in exploring self-deception from the perspective of philosophy of information, building from Arendt's foundational scholarship on deception to examine the twenty-first-century relationship between information and the self-deceived. Because Arendt suggests that self-deception can be widespread, stemming from a mendacity which alters the fabric of the real, it is important to establish what the real means in informational terms. To do so, I turn briefly to philosopher of information Luciano Floridi (2013), who popularised the term '*infosphere*' to denote '*the whole informational environment constituted by all informational entities (thus including information agents as well), their properties, interactions, processes, and mutual relations*' (Floridi, 2013, p. 6). Because information is so encompassing, the infosphere '*can also be used [as a term] synonymous with reality, or Being*' (Floridi, 2013, p. 6). The systems of mendacity which Arendt warns against can be understood as manipulations of the infosphere, manipulations which perhaps result in the same self-deception today that proliferated in the twentieth century. My research questions are thus:

1) What is informational self-deception?
2) What are the political and ethical stakes of informational self-deception?

To answer these research questions, I confront a central paradox inherent to self-deception as it is portrayed in philosophical literature: that between ontological and psychological renderings of the concept. Some theorists present self-deception ontologically, as a pervasive state of being-in-the-world. Put informationally, we exist as part of an imperfect and entropic infosphere, and so we are collectively subject to a certain level of self-deception. Other theorists, however, argue that self-deception is an active choice of a few rather than the passive condition of the many. Put informationally, some individuals are motivated towards self-deception, willingly ignoring or warping information that would rupture the image of the world they have crafted. In this paper, I label such active self-deception as psychological, as these renderings of self-deception often draw on more recent work in psychology to describe the partitioning necessary for such deceit. In this paper, I seek to stabilise the concept of self-deception within philosophy of information, examining informational self-deception as a concept which is maintained by the tension between an ontological condition of un-intentional deceit and a psychological condition of intentional deceit.

Method

My method for this paper involves confronting and deconstructing ontological configurations of self-deception and psychological configurations of self-deception. Is self-deception an ontological concept, the passive condition of humanity out of which we must actively work to escape? Or, rather, is self-deception a psychological concept, an active choice, something into which we fall rather than out of which we escape? In examining ontological self-deception and psychological self-deception, divergences in questions of self-deception's relationship to complacency, action, intentionality and motivation become apparent.

The subsequent deconstruction is explicitly Derridean. On a basic level, the project of Derridean deconstruction involves the examination, complication or creation of dichotomous and aporetic structure with the goal of creating productive, irresolvable tension (Derrida, 1992). Unlike a Hegelian dialectic, Derridean deconstruction does not strive for a mediating synthesis between a thesis and antithesis. Rather, it seeks to destabilise and maintain paradox, continually disturbing foundational assumptions in order to make ethical space for the yet-to-come. In this way, I do not seek to establish informational self-deception as a mediating synthesis between ontological and psychological renderings of the concept. Rather, I use deconstruction to argue that informational self-deception is itself maintained by the very aporias created between the ontological self-deception and psychological self-deception. Informational self-deception is active *and* passive, collective *and* individual, intentional *and* unintentional; it creates paradox, and in so doing it invokes the constant examination and destabilisation of the relationship between self, information and deception.

Self-deception as ontological

Ontological self-deception positions deception as a natural state of being-in-the-world; we exist naturally deceived, and it is our responsibility, therefore, to be critical, to scrutinise information and to work our way out of such deception. Arendt (1963), for example, seems to describe self-deception ontologically when she discusses entire citizenries as being complacent in a passive self-deception that perpetuated the harms of regimes. For philosophy of information, thinking of self-deception through the lens of an information ontology is interesting in that it explains how individuals who allow themselves to passively be swept into information silos fall prey to self-deception without actually *doing* anything. When we allow the infosphere to go unquestioned, we can become informationally self-deceived. Situating informational self-deception as an ontological concept is additionally beneficial in that it places the onus of responsibility to escape self-deception on each person regardless of any perceived allegiance to ideology, social group or bureaucracy. It inspires a moral charge to consistently question information gaps and our own positionality.

To establish ontological self-deception in discourse, I risk anachronism by turning to a pre-ontological thinker: Plato. Plato is considered pre-ontological because his

> *interest in the theory of the self, or egology, was ethico-political and epistemological, not yet ontological [...] It is mainly from Descartes onwards that the unity, identity, and continuity of the I, or self, as an entity become the subjects of an ontological investigation in their own right.* (Floridi, 2011, p. 552)

Though Plato's construction of the self necessarily varies from future iterations of the concept, his development of deception as a natural condition for humanity structured centuries of future theorising. Moreover, his theory of self-deception as tied to existence fits the contemporary use of the term *ontology* even if Plato himself came well before Descartes' *cogito*.

One of the clearest places within Plato's dialogues in which a theory of natural deception is outlined is in the allegory of the cave, which is found in book seven of his *Republic*. In this allegory, Plato

(1982/375BCE) describes humanity as prisoners chained in a cave, staring at shadows cast upon the cave wall, listening to echoes and believing them to be produced by the shadows. Having never experienced reality beyond the cave, we deceive ourselves in believing these shadows and echoes are reality. It is important to specify that Plato's allegoric prisoners do not begin in the illumination of knowledge only to be pushed into the cavern by some force or captor. Rather, they begin the allegory already chained to the cave, having only ever known the cave, falsely assuming the truth to be '*literally nothing but the shadows of images*' (Plato, 1982/375BCE, VII515). Wise prisoners can learn to shake their chains and drag themselves '*up a steep and rugged ascent*,' but this journey out of deception is painful, marked by disorientation and light that blinds (Plato, 1982/375BCE, VII515). It is an escape requiring effort: few can make the ascent, and those that do receive ire and vitriol from those prisoners still chained.

While steeped in figurative language, Plato's depiction of self-deception as a condition of existence in the *Republic* mirrors Socrates' own defence that he mounts for himself at his trial as found in Plato's (1998/399BCE) 'Apology of Socrates'. In this text, Socrates is standing trial, accused of corrupting the youth of Athens and of blasphemy towards the gods. Socrates claims that the real reason he is on trial is because he revealed to Athenians the deception in which they were complicit. He questions Athenians in high regard, the supposed experts in different crafts and fields, tearing away the veil of feigned knowledge to conclude in each case,

> *I am wiser than this man; it is likely that neither of us knows anything worthwhile, but he thinks he knows something when he does not, whereas when I do not know, neither do I think I know; so I am likely to be wiser than he to this small extent, that I do not think I know what I do not know.* (Plato, 1998/399BCE, 22d-e)

Those who purported to possess expertise were often uncritical of themselves and of their knowledge, and, when prompted by Socrates' questions, were left unable to account for the opinions they held. Therefore, the unsuccessful defence Socrates mounts for himself at his trial is that he did not participate in the self-deception of the masses. Self-deception as an ontological condition became a core tenant of Platonism.

Interestingly, Floridi's concept of the infosphere provides a novel way of approaching Plato's ancient allegory of the cave. Though we may not be born into a literal cave, we do exist both in and as a part of an imperfect, shadowy informational world. Just as the Platonic cave contained shadows and echoes, so too does the infosphere contain entropy. Entropy in information ethics is defined broadly as '*a quantity specifying the amount of disorder, degradation, or randomness in a system bearing energy or information*' (Floridi, 2013, p. 66). Insofar as entropy is informational disorder and degradation, entropy can cause deceit. This implicates a deception that operates on an ontological level, for the infosphere itself is an ontological concept. Put simply, we, as informational beings, participate in the informational Being of the infosphere, a Being containing the entropic potential of deception. Floridi (2013) argues, '*we are slowly accepting the idea that we are not Newtonian, standalone, unique entities, but rather informationally embodied organisms [...], mutually connected and embedded in an informational environment, the infosphere*' (p. 14). There is currently '*a re-conceptualization of our ontology in informational terms*' (Floridi, 2013, p. 10). Therefore, ontological self-deception can be attributed to the fact that entropy exists in the infosphere, the informational reality in which we are ontologically imbedded. Because of this, we must be continually weary of the semantic content we receive, the information caves in which we find ourselves. Succumbing to self-deception may be conceived as a moral harm within the macro-ethical framework of information ethics insofar as it perpetuates or creates entropy within the infosphere.

Floridi's recasting of ontology along informational lines also implicates ICTs as re-ontologising technologies '*because they create and re-engineer whole realities that the user is then enabled to*

inhabit' (Flordi, 2013, p. 97). How can ICTs, as re-ontologising technologies, contribute to or dispel ontological self-deception? This question connects to philosophy of information research on the creation and proliferation of conspiracy theories in the twenty-first century, theories which thrive on the collective and participatory environments engendered by ICTs (Sicart, 2018; Zeeuw and Gekker, 2023). Research at the intersection of game theory and media theory positions conspiracy theories as most impactful when they are able to successfully engage a diffuse collective to participate in world-building practices; for example, '*conspiratorial metanarratives like QAnon construct an engrossing and immersive communal world for its adherents, cultivating (or strengthening) belief (regardless of original intent) and shaping (political) behaviors*' (Zeeuw and Gekker, 2023, p. 10). Beyond static *theory*, contemporary conspiracy movements invoke the re-ontologising nature of ICTs to fiction an (un)reality which is driven by '*engagement, fuelled by fannish "fictioning" practices of those in the know, and aided by the hyperstitional unbelief and mythos of its followers [...] building alternate realities while disrupting existing ones*' (Zeeuw and Gekker, 2023, p. 5). This ability to build alternate reality, to superimpose the phantasmic over the ordinary, implicates a self-deception operating at an ontological level, using the re-ontologising nature of ICTs to build entropic information caves in which being-in-the-world is redefined.

In this section, I presented arguments for understanding self-deception ontologically; it is a diffuse condition of the many which echoes throughout the history of philosophy. Socrates observed in Athenians a collective and participatory self-deception woven into the very fabric of being-in-the-world. Such ontological self-deception bred a complacency which served those in power, and, when Socrates began to draw forward and question such a collective deception, he was sentenced to death. Centuries later, Arendt (1963) also wrote with horror on a collective self-deception that defined being-in-the-the-Third-Reich. This self-deception likewise encouraged a complacency which served those in power and resulted in massacre. Now, as the twenty-first century progresses and as being-in-the-world is understood in increasingly information ways, there seems a moral imperative to once again recognise the collective and ontological nature of self-deception, to search for entropic shadows in the infosphere and to question the role of ICTs in casting or dispelling these shadows.

Self-deception as psychological

The second way to construe self-deception is as a psychological concept. In this understanding, the motivations and intentions of individuals are centred. Psychological self-deception is conceived as an active and motivated choice of a select few rather than as a passive and ontological condition of the many. Because it is tied to intention and motivation, it is usually, though not without debate, separated in scholarship from unintentional and uncontrollable delusions or psychotic breaks (Bayne and Fernández, 2009). Psychological self-deception holds allure for philosophy of information because it opens questions surrounding the motivations of self-deception and whether these motivations are more prevalent in specific individuals. Thomas Froelich (2017, 2019, 2021), for example, uses psychological source material to examine self-deception among individuals who support Trump, engaging concepts of intentionality and responsibility in relation to the consumption and production of semantic content. In this section, I present a brief history of psychological self-deception in philosophical literature before connecting such self-deception to concepts of information behaviour and to today's political environment.

Just as I turned to a pre-ontological thinker to establish ontological self-deception, so too do I turn to a pre-psychological thinker to root psychological self-deception; I first examine how René Descartes reversed Platonic assumptions, before establishing more contemporary strands of the theory. Of course, it would be anachronistic to assume that Descartes, a seventeenth century thinker, was himself interested in psychology, a discipline which emerged in the nineteenth century. However, I argue that self-deception conceptually transformed from a collective to an

individual concept through Descartes, and this transformation paved the way for future psychological theorising.

When Descartes was writing in the seventeenth century, most of his contemporaries ascribed to Plato's rendering of self-deception as a pervasive state of being-in-the-world. Blaise Pascal (1995/1670), for example, lived and wrote in France at the same time as Descartes, and he characterised persons as '*nothing but a subject full of natural error [...] Nothing shows him the truth, everything deceives him*' (p. 12). Both reason and the senses '*compete in lies and deceptions*' (Pascal, 1995/1670, p. 13). In other words, we are Platonic prisoners existing naturally in a cave of deceit.

Descartes, however, viewed self-deception as a barrier to the pursuit of knowledge through reason. Therefore, he begins his *Meditations on a First Philosophy* by taking seriously the question of self-deception:

> *as I sometimes imagine that others deceive themselves in the things which they think they know best, how do I know that I am not deceived every time that I add two and three, or count the sides of a square, or judge of things yet simpler, if anything simpler can be imagined?* (Descartes, 1911/1641, p. 8)

Here Descartes is interested in questioning if reason, even the most basic reasoning of arithmetic and geometry, leads to anything more substantial than ever-changing shadows of truth. If we are deceived even in this most simple thought, then Descartes (1911/1641) reasons that such a deceived state of being must result from a deception operating on a cosmic scale. He pushes this logic to its limit in his great project of doubt:

> *I shall then suppose, not that God who is supremely good and the fountain of truth, but some evil genius not less powerful than deceitful, has employed his whole energies in deceiving me; I shall consider that the heavens, the earth, colours, figures, sound, and all other external things are nought but the illusions and dreams of which this genius has availed himself in order to lay traps for my credulity.* (p.8)

Descartes' starting point is therefore consistent with a Platonic level of self-deception. What if we only see shadows? What if we are born into deception?

Through his meditations Descartes leads the reader to a quiet conclusion, often overlooked in favour of his *cogito*: subjectivity does not, after all, exist in a natural state of self-deception. This he proves by arguing that God exists and is supremely good; '*From this it is manifest that He cannot be a deceiver, since the light of nature teaches us that fraud and deception necessarily proceed from some defect*' (Descartes, 1911/1641, p. 19). In other words, God would not chain us to a cave. Descartes (1911/1641) thus ends his *Meditations* by dismissing the question of self-deception: '*I ought no longer to fear that falsity may be found in matters every day presented to me by my senses [...] For because God is in no wise a deceiver, it follows that I am not deceived*' (p. 32). In this way, Descartes quietly flipped a centuries-old philosophical tradition, and self-deception became something to fall into through error rather than to escape from through effort. Whereas Platonists assumed a certain amount of unknowability inherent to the world tied to ontological deception, Cartesians assumed an amount of knowability inherent to the world tied to a trust that such ontological deception is illusory. Though Christian theological contemporaries of Descartes disagreed with him regarding the nature of God excluding the possibility of widespread deception (see, for example, Pascal, 1995/1675), the Cartesian reversal of ontological self-deception took root, and self-deception became an indicator of individual failings rather than a natural state of being in-the-world.

After psychology's emergence as a discipline in the nineteenth century, self-deception as an action tied to individuals became something that could be studied through the lens of motivations and

intentions. Doing so, however, first involves confronting another paradox inherent to the concept; if to lie requires that one both knows the truth and distorts it with the intention to deceive, then to lie to oneself requires that one both knows the truth and intentionally distorts it to instead believe a falsity – that one simultaneously knows *p* but believes ~*p* (Baghramian and Nicholson, 2013). Can the deceiver and the deceived coalesce within an individual in this way?

Across both philosophy and psychology there are sceptics who reject the possibility self-deception based on this paradox, arguing that the concept is a logical impossibility better explained by related phenomenon like illusion, ideology or error (Baghramian and Nicholson, 2013). Derrida (2002, p. 67), for example, expresses doubt over the possibility self-deception, arguing that to lie to oneself is '*logically incompatible with the rigor of any classical concept of the lie [...]. To lie will always mean to deceive the other intentionally [...], while knowing what it is that one is deliberately hiding, therefore while not lying to oneself*'. Despite such criticism over the paradoxical convergence of deceiver and deceived, other theorists overcome the contradictory nature of psychological self-deception by relying on the concept of mental partitioning '*to explain how a self-deceiver holds contradictory beliefs simultaneously or brings about a false belief intentionally*' (Baghramian and Nicholson, 2013, p. 1020). Partitioning, which has origins in Freud's division of consciousness, is the theory that the mind is itself subdivided to an extent that '*the self-deceiver's beliefs can be contradictory but concurrent*' (Baghramian and Nicholson, 2013, p. 1020). Arendt interestingly discusses such partitioning within the self from a philosophical rather than psychological perspective. She argues that the self is not some unbroken unity. Rather, the self is divided and can also be placed in relation to itself; I can question my own choices, hear my own thoughts, ponder my own decisions, divide myself so as to recognise the difference within myself. This is what leads Arendt (1971) to conclude:

> *Human consciousness suggests that difference and otherness, which are such outstanding characteristics of the world of appearances as it is given to man as his habitat among a plurality of things, are the very conditions for the existence of man's ego as well* (p. 442).

Partitioning does not inherently pose a threat to selfhood, and it does not have to be understood from within the pages of psychology scholarship. Rather, such fragmentation is woven into the fabric of selfhood and is studied philosophically as well psychologically.

Partitioning in and of itself is not an indicator of self-deception. In fact, the ability to question one's own biases, beliefs or knowledge relies on partitioning and is an important way in which one can escape self-deception (Arendt, 1971). However, if the mental self is not a cohesive unity but instead a composite of parts, then perhaps there is no contradiction between the coalescing deceiver and deceived: '*There are a variety of dissociations between seemingly continuous mental processes that ensure that the mental processes that are the target of self-deception do not have access to the same information as the mental processes deceiving the self*' (von Hippel and Trivers, 2011, p. 6). Partitioning strategies assert that self-deception is not the logical impossibility that critics assume.

I suggest that such psychological self-deception within the partitioned self is of particular interest to information science and philosophy of information, as work at the intersection of psychology, behavioural science and cognitive science indicates that the self-deceptive potential of partitioning is tied to information processing. Psychologist William von Hippel and evolutionary biologist Robert Trivers (2011, p. 1) argue that '*what marks all [...] varieties of self-deception is that people favor welcome over unwelcome information in a manner that reflects their goals or motivations*'. In other words, psychological self-deception stems from motivated preference for information that confirms rather than confronts biases. While one may unconsciously or subconsciously recognise inaccurate or biased information processing, one's consciousness is willingly deceived into accepting low quality or biased information over quality or unbiased information (von Hippel and Trivers, 2011).

This rendering of individual, psychological self-deception is similar to recent work in information science by Reijo Savolainen (2022) on motivated reasoning. Motivated reasoning, proposed by Ziva Kunda in 1990, is a psychological theory which suggests that '*motivation can be construed as affecting the process of reasoning: forming impressions, determining one's beliefs and attitudes, evaluating evidence, and making decisions*' (Kunda, 1990, p. 480). It is connected to self-deception by scholars who see it as a path to self-deception, an explanation for the motivations behind becoming self-deceived (von Hippel and Trivers, 2011). Savolainen (2022) argues that motivated reasoning provides a fruitful framework for understanding why individuals accepted health-related misinformation during the Covid pandemic even when that misinformation had sometimes deadly consequences. My present examination of self-deception can serve to corroborate this line of research, building out a philosophical foundation to support future work on the stakes of such motivated deception.

As opposed to motivated deception, some work at the intersection of psychology and information science suggests that the acceptance of political misinformation '*may originate from a lack of analytical thinking, rather than the tendency to adhere to politically concordant information content supporting an individual's partisan beliefs*' (Savolainen, 2022). Psychologists Gordon Pennycook et al. (2020), for example, attribute individual acceptance of Covid-related misinformation to an absence of thought rather than to motivated reasoning, arguing that people accept information without taking the additional steps necessary to evaluate or authenticate that information. This line of reasoning is similar to Arendt's (1971, p. 417) own description of self-deception in relation to Eichmann; '*the only specific characteristic one could detect in his past as well as in his behavior during the trial and the preceding police examination was something entirely negative: it was not stupidity but a curious, quite authentic inability to think*'. A lack of thought is thus associated with accepting untruths. Can we assume, however, that this rendering of self-deception as the absence of thought is, in fact, divorced from motivation or intentionality? Could motivated reasoning and the motivated absence-of-reasoning be two similar, informational paths to self-deception? If so, what could motivate the refusal to engage in thought?

To answer these questions, I return to von Haben and Trivers (2011), who note distinct evolutionary advantages of self-deception, advantages which apply both to motivated reasoning and to a motivated lack-of-reasoning. The first advantage of self-deception '*is that by deceiving themselves, people can better deceive others, because they no longer emit the cues of consciously mediated deception that could reveal their deceptive intent*' (p. 4). This is an advantage also recognised by Arendt (1968) who attributes the modern lie to the proliferation of self-deceived deceivers. Today, we might similarly question if figures in power are motivated to self-deception in order to propagate systemic mendacity. In this way, the psychological self-deception of individuals may engender ontological self-deception among a citizenry though manipulating our entropic infosphere. The second evolutionary advantage of self-deception '*is that by deceiving themselves, people are able to avoid the cognitive costs of consciously mediated deception*' (von Haben and Trivers, 2011, p. 4). Again, I argue that this motivation to self-deception is seen in Arendt's (1968) analysis, specifically when she notes in Eichmann an un-critical complacency and lack of remorse when faced with his crimes in his post-war trial. Perhaps this self-deceived lack of thought stems from the motivation to avoid the cognitive costs of accepting moral responsibility for the systemic murder of millions. While this is an extreme example, philosophers of information today should likewise question the cognitive costs at stake in instances of psychological self-deception, whether such self-deception is a result of motivated reasoning or a motivated lack-of-reasoning.

In this section, I explored how self-deception transformed in philosophical literature from a passive condition of the many to an active shortcoming of the individual, a transformation which laid the foundation for the development of psychological self-deception. Psychological self-deception relies on partitioning in order to explain how one might know the truth and then

intentionally distort it in order to believe a falsity. Such partitioning and the self-deception it engenders connects to questions of information behaviour and motivated reasoning, concepts which are used to understand why one might accept misinformation in the face of quality information. I argue that a motivated lack-of-reasoning can also lead to self-deception; motivations to avoid critical thought and fall into deceit can include the desire to deceive more effectively and the desire to avoid cognitive load. Ultimately, psychological self-deception is beneficial for philosophy of information in raising questions of individual information behaviour and motivation, confronting the reasons for, variances in and costs of such deceit.

A concluding reconstruction: informational self-deception as ontological and psychological

This deconstruction thus creates an irresolvable tension between self-deception as ontological, passive and pervasive, and self-deception as psychological, motivated and individual, with both renderings of the concept contributing unique and beneficial viewpoints to philosophy of information. Which, then, is the correct path to take when studying such deception from an informational perspective? The answer, I argue, is both: informational self-deception is maintained by the aporetic structure of the two halves of this deconstruction.

In a way, this is the answer Arendt herself seems to assume. Though she never separates or labels these two different depictions of self-deception, she allows both ontological and psychological renderings of self-deception to entangle in her writing. For example, she describes '*the practice of self-deception*' as '*so widespread*' in Germany during the Nazi regime that it had become '*almost a moral prerequisite for survival*' (Arendt, 1963). In this way, information manipulation can lead to a '*complete rearrangement of the whole factual texture—the making of another reality [...] into which [modern lies...] fit without seam, [...] as the facts fitted into their own original context*' (Arendt, 1968, p. 253). Such passive and ontological self-deception bred a complacency that perpetuated extreme harms in the twentieth century. In other cases, however, Arendt (1968, p. 254) uses self-deception to call out specific governmental bureaucrats who willingly '*fall prey to [their] own fabrications*' in order '*to create a semblance of truthfulness*' for their entangled deceptions. In these areas of scholarship, Arendt assumes more of a psychological perspective towards self-deception, exploring specific motivations and consequences for falling prey to one's own deceit. I argue that this tension implicit in Arendt's writing opens space to explore informational self-deception in her work and case studies and to draw these theories into the diffuse systems of mendacity imbedded in today's infosphere.

Ontological self-deception is supported by Floridi's ontological portrayal of the infosphere as Being. Insofar as we, as beings, participate in an entropic infosphere, as Being, we are subject to a basic level of ontological self-deception that we must consistently strive to escape. Ontological self-deception places the responsibility to be critical of received semantic content on each individual regardless of ideology or political orientation. It includes the Socratic call to serve as a gadfly, continually questioning oneself and others. Additionally, ontological self-deception provides a philosophical explanation for the successes of contemporary conspiracy movements which use world-building strategies and the re-ontologising nature of ICTs to build alternative realities. The ethical and political stakes of fostering deception and encouraging complacency are merely introduced here, and a foundation is set for future, more in depth exploration of this topic. Ultimately, however, ontological self-deception does not explain variance in levels of self-deception, and it leaves out important questions surrounding intentionality, motivation and partitioning. For this, psychological self-deception is important. Rather than focus on how the many are unintentionally self-deceived, psychological self-deception focuses on the way in which specific individuals become intentionally self-deceived. This connects to existent work in information science on motivated reasoning while simultaneously introducing a motivated lack-of-reasoning as an additional path to psychological self-deception. Examining the relationship

between psychological self-deception, motivated reasoning and motivated lack-of-reasoning implicates those who propagate deceptions as well as those who succumb to deceptions, providing a foundation in philosophy of information for additional theorising on specific instances of self-deception.

Therefore, this deconstruction positions informational self-deception as, at the same time, ontological and psychological, the unintentional condition of the many and the intentional condition of the few. Though self-deception is a concept which haunts centuries of philosophical literature, philosophy of information is positioned to offer a new perspective on such mendacity, engaging the twenty-first century relationship between individuals, communities and our entropic infosphere.

Acknowledgements

I would like to thank Marija Dalbello for advising this research project. Her willingness to review drafts and provide insightful feedback greatly improved the present paper. I would also like to extend my gratitude to my reviewers for their careful reading of my work.

About the author

Juliana Mestre is a PhD candidate in Communication, Information, and Media at Rutgers University. Her research engages philosophy of information from a post-structural and deconstructive perspective. Juliana.mestre@rutgers.edu

References

Arendt, H. (1963, February 8). Eichmann in Jerusalem—I. *The New Yorker*. https://web.archive.org/web/20250309023545/https://www.newyorker.com/magazine/1963/02/16/eichmann-in-jerusalem-i

Arendt, H. (1968). *Between past and future: eight exercises in political thought* (Engl. ed.). Viking Press.

Arendt, H. (1971). Thinking and moral considerations: a lecture. *Social Research*, 38(3), 417–446.

Arendt, H. (1972). *Crises of the republic: lying in politics, civil disobedience on violence, thoughts on politics, and revolution*. Harcourt Brace Jovanovich.

Arendt, H. (2006). *Eichmann in Jerusalem a report on the banality of evil*. Penguin Books.

Baghramian, M., & Nicholson, A. (2013). The puzzle of self-deception. *Philosophy Compass*, 8(11), 1018–1029. https://doi.org/10.1111/phc3.12083

Bayne, T., & Fernández, J. (2009). Delusion and self-deception: mapping the terrain. In T. Bayne & J. Fernández (Eds.), *Delusion and self-deception: affective and motivational influences on belief formation* (pp. 1-21). Routledge.

Derrida, J. (1992). Force of law: the 'mystical foundation of authority'. In D. Cornell, M. Rosenfeld, & D. G. Carlson (Eds.), *Deconstruction and the possibility of justice* (pp. 1–67). Routledge.

Derrida, J. (2002). *Without alibi* (P. Kamuf, Ed.). Stanford University Press.

Descartes, R. (1911). Meditations on first philosophy. In E. Haldane (Trans.), *The philosophical works of Descartes*. Cambridge University Press. Originally published in 1641.

Deweese-Boyd, I. (2023). Self-deception. *The Stanford Encyclopedia of Philosophy*. https://archive.ph/BeRG

Fallis, D. (2010). Lying and deception. *Philosophers' Imprint*, 10(11), 1-22.

Fallis, D. (2011). Floridi on disinformation. *Etica & Politica*, 13(2), 201–214.

Fallis, D. (2015). What is disinformation? *Library Trends*, 63(3), 401–426. https://doi.org/10.1353/lib.2015.0014

Fallis, D. (2016). Mis- and Dis- Information. In L. Floridi (Ed.), *The Routledge handbook of philosophy of information* (pp. 332–346). Routledge.

Fallis, D. (2022). Deceiving versus manipulating: an evidence-based definition of deception. *Analytic Philosophy*, 1–18. https://doi.org/10.1111/phib.12282

Floridi, L. (2011). The informational nature of personal identity. *Minds and Machines*, *21*(4), 549–566. https://doi.org/10.1007/s11023-011-9259-6

Floridi, L. (2013). *The ethics of information*. Oxford University Press.

Floridi, L., & Illari, P. (Eds.). (2014). *The philosophy of information quality*. Springer International Publishing . https://doi.org/10.1007/978-3-319-07121-3

Froehlich, T. J. (2017). A not-so-brief account of current information ethics: the ethics of ignorance, missing information, misinformation, disinformation and other forms of deception or incompetence. *BiD: Textos Universitaris de Biblioteconomia i Documentació*, 39, 1–14. https://web.archive.org/web/20250328200157/https://bid.ub.edu/en/39/froehlich.htm

Froehlich, T. J. (2019). The role of pseudo-cognitive authorities and self-deception in the dissemination of fake news. *Open Information Science*, 3(1), 115–136. https://doi.org/10.1515/opis-2019-0009

Froehlich, T., J. (2021). A disinformation-misinformation ecology: the case of Trump. In J. Višňovský & J. Radošinská (Eds.), *Fake news is bad news—hoaxes, half-truths and the nature of today's journalism*. IntechOpen. https://doi.org/10.5772/intechopen.95000

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

La Caze, M. (2017). It's easier to lie if you believe it yourself: Derrida, Arendt, and the modern lie. *Law, Culture and the Humanities*, 13(2), 193–210. https://doi.org/10.1177/1743872113485032

Palomo, M. (2021). How disinformation kills: philosophical challenges in the post-Covid society. *History and Philosophy of the Life Sciences*, 43(2), 51. https://doi.org/10.1007/s40656-021-00408-4

Pascal, B. (1995). *Pensées* (A. J. Krailsheimer, Trans.; Rev. ed). Penguin Books. Originally published in 1670.

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, 31(7), 770–780. https://doi.org/10.1177/0956797620939054

Plato. (1982). *The Republic* (B. Jowett, Trans.). Random House. Originally published in 375BCE.

Plato. (1998). Apology of Socrates. In T. G. West & G. S. West (Trans.), *Four texts on Socrates: Plato's euthyphro, apology, and crito, and Aristophanes' clouds* (Rev. ed, pp. 16–23). Cornell University Press. Originally published in 399BCE.

Savolainen, R. (2022). What drives people to prefer health-related misinformation? The viewpoint of motivated reasoning. *Information Research: An International Electronic Journal*, *27*(2). https://doi.org/10.47989/irpaper927

Sicart, M. (2018). Play in the information age. *Philosophy & Technology*, *32*(3), 517–534. https://doi.org/10.1007/s13347-018-0311-9

Søe, S. O. (2018). Algorithmic detection of misinformation and disinformation: Gricean perspectives. *Journal of Documentation*, *74*(2), 309–332. https://doi.org/10.1108/JD-05-2017-0075

Søe, S. O. (2021). A unified account of information, misinformation, and disinformation. *Synthese*, *198*(6), 5929–5949. https://doi.org/10.1007/s11229-019-02444-x

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

Zeeuw, D. D., & Gekker, A. (2023). A god-tier LARP? QAnon as conspiracy fictioning. *Social Media + Society*, *9*(1), 20563051231157300. https://doi.org/10.1177/20563051231157300