

## Bullshit as a Practical Strategy for Self-Deceptive Narrators

Leslie A. Howe

**ABSTRACT.** This paper argues that bullshit is a practical resource for self-deceiving individuals, or those who merely prefer to avoid self-examination, insofar as it is able to provide a mask for poor doxastic hygiene. While self-deception and bullshit are distinct phenomena, and bullshit does not cause self-deception, bullshit disrupts the capacity to interrogate the motivational biases that fuel deception. The communicative misdirection engaged in by ordinary social bullshitters is applied reflexively by the self-deceiver to distort, evade, and obfuscate the self-deceiver's self-accounting. This discussion presupposes a broadly narrative approach to self-awareness and discusses how a motivated susceptibility to bullshit offers an explanation how our reports about ourselves can be frequently at odds with reality, and suggests that a complacency about rational validation of belief outside of self-certainty, i.e., the prevalence of bullshit, is even more of a threat than Frankfurt thought it to be.

Philosophical discussions about bullshit have largely followed Harry Frankfurt's (1998) lead in concentrating on the bullshitter's relationship to the *truth* of what is said and the bullshitter's lack of regard for the same. Philosophical analyses of self-deception have likewise tended to focus on the self-deceiver's doxastic commitments and failures. Given the widespread use of bullshit to mislead about the truth, one might conclude two things: that both bullshit and self-deception are about inducing false beliefs and that bullshit might be a useful method for self-deceivers. If bullshit can be effective in obscuring one's cognitive commitments to others, why not use it to similarly obscure oneself from oneself?

While bullshit may have distinct instrumental value for the self-deceiver, it is neither a means *to* self-deception by itself nor a reliable indicator of it. Rather, I argue that if one is *already* self-deceived, bullshit may help to maintain that condition, just as it may serve to obscure a subject's other, non-self-deceptive, characteristics. Given my purposes here, then, I will not be offering an extensive analysis of self-deception as such. I assume that it is a condition for which humans have a predilection, but its specific analysis should have limited relevance to the question of whether one can bullshit oneself about one's self-deceit.

In the following, I give a brief and very basic summation of the structure of self-deceptive practice. Note that the question under examination here is not strictly that of how self-deception itself, as a given doxastic state, should be analysed rather than how the self-deceiver, or the person who fears that they may be self-deceived but would rather not know, responds to this background condition. There is a truth at stake here, but the question at hand is about behaviour rather than belief, i.e., given a doxastic conflict (or a reluctance to consider the possibility of such a conflict), what practical failures of action in relation to belief might

ensue. I argue that bullshitting oneself is a possible response to such a situation. I give then a more detailed discussion of some recent discussion on bullshit, focussing on its strategic deployment for purposes of evasion and obfuscation. I then return to a consideration of the close connection between self-deception and bullshit through the characterisation of self-deception in terms of pretence. These two discussions highlight the circumstance that, as bullshit is a communicative action, self-bullshit is a narrative practice. The validation of self-narrative as fundamentally truthful, then, needs to include close attention to the incentives for self-narrators to bullshit their way through their self-accounting.

### ***Truth and Preferred Realities***

It is unlikely that anyone *sets out* to hold incompatible beliefs in the full recognition of their incompatibility.<sup>1</sup> People evidently do seem to *end up* doing so, but the sort of ordinary self-deception the present discussion assumes is that where *S* has reason to believe *p* but also has what *S* takes to be reason to weigh  $\sim p$  as more likely, and consequently to believe  $\sim p$ , where we also think that *S* ought to know better than to do so. In such cases, both *p* and  $\sim p$  have epistemic force for *S*, but  $\sim p$  is given more weight because *S* allows practical concerns to override normal (neutral) epistemic controls on belief. This needn't suppose either cognitive incompetence or that practical reasons lack doxastic referent, but rather that a heavier weighting is given to a desire that something *be* true, on top of whatever evidence might be available for  $\sim p$ , over an epistemic commitment to, and practice of checking whether,  $\sim p$  really is true. It is in this sense that we would say that *S* is motivationally biased in favour of  $\sim p$  despite inferior evidence, *S*'s desire for  $\sim p$  to be true outweighing even what might be superior grounds for *p*, of which *S* is properly aware.<sup>2</sup> It is the desire for something to be true rather than something else that tips us into self-deception and gives us a reason to distort the evidence.<sup>3</sup> Sartre's examples of bad faith exemplify this situation as well: the woman and the waiter in the café both have sufficient evidence for their materiality/transcendence but both have motivated reasons for ignoring these in certain circumstances and for certain periods of time (Sartre 1966, 86-96).<sup>4</sup> Insofar as these practical concerns are not always in the forefront of

---

<sup>1</sup> I only say "unlikely" here as I take religious faith of a Kierkegaardian sort to (apparently) involve such a proposition.

<sup>2</sup> I (mostly) follow Mele (2001) here.

<sup>3</sup> It is because there is a desire that something be true, or that something (relatively) specific be obtained, that "desire" here is meant to indicate a reason for acting, and not something as nonspecific as appetite, which I take to have much more limited epistemic content.

<sup>4</sup> Gibson (2020) captures this situation exceptionally well in terms of intentional failure (omission) rather than intentional action. What needs then to be explained is how (what

attention, self-deception's persistence may be hidden by an episodic rather than constantly present character: in circumstances **C** one acknowledges **p** but in **D** one acts as if  $\sim p$  is true. Insofar as the self-deceiver is able to mask the inconsistency in this way (as with a little made-to-measure bullshit) it is more difficult to eradicate.

What is of particular interest here is the motivation for self-deception rather than the doxastic state that is its realisation. What could make this cognitive wreckage worthwhile? The same question is just as or more pertinent when we look at the masking of self-deception, for which self-bullshit is particularly well suited. Bullshitting oneself about self-deception does not require that one actually be self-deceived. One may simply be concerned that one could be self-deceived and so resolve, perversely, to make purposeful self-examination as difficult as possible, or one might just not care one way or the other ("what is truth, anyway, but an oppressive social-linguistic concept?"). Bullshitting oneself about what one believes and self-deception are distinct phenomena. The aspiring self-deceiver must first have a specific doxastic aim (the belief that  $\sim p$ ), which is motivated by whatever desire it is that requires that  $\sim p$  be actually true. But, following Frankfurt, the bullshitter has no such allegiance to truth, and so self-deception masking bullshit does not work by defending definite if false claims but by making such claims incoherent or irrelevant; it misdirects, evades, and obscures, causing a kind of cognitive havoc. Unlike what goes on with self-deception, bullshit is no respecter of propositional specificity; it is not a state but an action, a narrative performance intended to *do* something. All this means that we cannot say that one bullshits oneself in order to, or as a means of, self-deceiving. One could do either without doing the other. But bullshit, as we shall see further on, can keep one in thrall to self-deception because it makes it so difficult to see one's way out of it, whether one bullshits oneself deliberately or is "merely" enmeshed in a social culture of bullshit.

Self-targetted bullshit, nevertheless, may serve self-deception, not by causing it (necessarily), but principally by short-circuiting attempts at self-examination and doxastic review that would uproot self-deception. In effect, if self-deception is a first-order condition, self-targetted or implicated bullshit is a second-order device for making sure that self-deception is either not resolved or not discovered. It is a preservative measure that works by deflecting or bamboozling the subject's attention to themselves, just as other-directed bullshit operates to deflect others from the speaker's intentions and commitments.

### ***Avoiding and distorting reality***

Unlike self-deception, bullshit is generally taken to be a communicative and paradigmatically linguistic phenomenon. It should be said at the outset that ordinary English

---

mechanisms the agent resorts to) in order to maintain the contrived nescience of competing evidence.

usage applies the term “bullshit” to a wide range of linguistic gambits, as well as actions<sup>5</sup>, the common characteristic being that all are judged derisive. Sometimes bullshit is carefully crafted meaninglessness (as per Cohen (2002)) and sometimes it is incoherent self-contradictory rambling that does little more than consume unrecoverable time. “Bullshit” is a colloquial term used to express contempt for a range of speech types deployed in a range of situations, but with specific aims.

Frankfurt objects to the bullshitter on two grounds: that he does not care for the truth and because he is dishonest about “his enterprise” (Frankfurt, 130). It is the bullshitter’s lack of regard for the truth, including whether his audience ends up believing either truth or falsehood, that is the basis for Frankfurt’s condemnation of bullshit as “a greater enemy of the truth than lies are” (Frankfurt, 132). There may well be some bullshitters who have this as an explicit goal, but the vast majority of bullshit is spread on smaller fields. The primary concern of this paper is how we end up bullshitting ourselves, but first we need to understand what the average intersubjective bullshitter aims to accomplish and why their intentions need to be kept covert. This also means that, although we have good reason to be careful about what counts as bullshit (the product), the focus in the following is on the bullshitter (the producer of bullshit).

### *Misleading through Performative Sincerity*

Bullshit is a type of misleading communication, usually in the form of speech. The audience is meant to draw certain conclusions from it, ones that are *in some way* inaccurate, whether about the subject of the speech or about the speaker. For example, a university or corporation issues a statement pledging itself to some kind of future behaviour; the bullshit may be in the substance of the pledge and/or in the expressed sincerity of the official delivering it. Bullshit differs from lying (following Frankfurt) in that the bullshitter, unlike the liar, is not necessarily attempting to get the audience to believe any specific proposition that is the explicit content of the speech. That is, in response to a direct question, the bullshitter may utter a great many propositions, some of which may be true and some false, in order to extract from the questioner either a positive response or a simple cessation of questioning. The bullshitter presumably doesn’t care whether anything true has been said, but only that the questioner is satisfied, or just stops asking questions. A really accomplished bullshitter will also win approval from the questioner (e.g., “he *is* bright, isn’t he”). *This* is the aim of the exercise, just as we are meant to see the multinational corporation or university administration as ethical and progressive. Thus, what is said is bullshit insofar as it is epistemically insincere (the speaker has no epistemic commitment to it) and it is also bullshit insofar as the purpose in uttering it is not to advance truth or justice but to evade heat or acquire praise.

---

<sup>5</sup> As a sport spectator might say of a referee’s call, “that was bullshit” or “that was a bullshit penalty”, the meaning of which is quite a bit more complicated than just “that was not true”, but that there was something about the interaction and its motivations that is suspect.

Whether the object of the misdirection is the substantive content of speech or the attitudes of the speaker, what is common to both is the hiding of it. Or so it is in most cases. It is true that sometimes we know we are being bullshitted and the bullshitter knows that we know and could care less that we know. There is a social cost to calling out bullshit in certain circumstances—sometimes social harmony leads us to bite our tongues and sometimes it is a significant power imbalance that makes it pointless or dangerous to say what we see. Indeed, as Jonathan Lear points out, being able to bullshit with impunity is itself a demonstration of power (Lear 2005; Borman 2011, 130-132). Nevertheless, even as bullshit becomes more common and more brazen, it remains a practice a characteristic part of which is its denial of itself. Even when we suspect bullshit but are restrained by politeness from saying so, successful bullshit contains its own deniability, if only in its amorphousness. Who can be sure exactly what the bullshitter just said, anyway? Or, if the bullshit is not of the unclarifiably obscure sort, the deflection of attention from the bullshitter's own shortcomings in the truthfulness department to some other attention seizing issue prevents the audience from examining the speaker's credentials more closely. These are the two central tactics of bullshit: evasion and obfuscation.

### *Evasion*

What is standardly evaded by the bullshitter is truth. A number of commentators have argued against Frankfurt's claim that the bullshitter is dangerous because of his lack of regard for the truth that, in fact, the bullshitter *must* have at least some concern for it, if only instrumentally.<sup>6</sup> There is a certain amount of confusion here, however, mainly because of a lack of sufficient distinction between aims, on the one hand, and strategies and tactics, on the other. The aims of bullshit can be many but generally concern the gaining or maintaining of status and saving face, or, sometimes, mitigating harm to others (leaving aside more pathological motivations, such as malice). Context provides the specific aim but, in many cases, bullshitters bullshit in order to convince others that they are smarter, kinder, more *au courant*, and just generally *better* than they are or could be bothered to work to be. This includes mitigating behaviour, such as when someone doesn't have the answer to a question, knows they are going to look bad, but bullshits a response in the hope that they will look *less* bad. Much bullshit that is produced in such circumstances is simply an attempt at deflection—a more sophisticated version of “look, a squirrel!”. Similarly, one might bullshit in order to stall for time, i.e., one fills the conversational air with waffle until one can think up something more substantive or defensible. Thus, the truth that is evaded (the aim) is one about the bullshitter: that they don't know, don't care, are trying to convince you of something that they don't really want to specify, don't want to say something that will hurt the listener or make the audience dislike them, and so on. How they go about the overall strategy of evasion (the tactics) may or may not involve

---

<sup>6</sup> E.g., Sara Bernall, (2006), Scott Kimbrough, (2006). Also, see Thomas L. Carson (2009), 181-6 who is much more careful and only points out that the bullshitter can be concerned with truth.

telling the truth, lying, misleading,<sup>7</sup> making things up, or talking about a completely different subject.

Bullshit is rarely about the specific propositional content of the speech that is used by the bullshitter. Hence, any claim that bullshit is necessarily connected in some way with truth should be carefully weighed. Bernall, for example, suggests that bullshit is truth-connected because of the desire to obscure part of it, which is done either through the indirect implication of falsehood, or through distraction by offering “evidence” for *p* that is other than the evidence on which the speaker actually relies (Bernall (2006), 64-66).<sup>8</sup> This overstates the bullshitter’s implicative requirements. Bullshit frequently works by simply overwhelming the attention of the audience. Any old bullshit will do. There is no need to imply a falsehood; only to say something *else*. One can bullshit by saying true, though irrelevant things, such as by unleashing a barrage of pseudo-erudition or actual fact in order to demonstrate either illusory or real expertise in order to distract from the critical truth that the speaker lacks the *relevant* knowledge that is immediately at issue. Moreover, insofar as bullshit is engaged in as a cover for ignorance, in particular, the perpetrator can have no truth-reference from which to operate—other than the truth that they are ignorant. Additionally, insofar as the speaker is also self-deceived about this, the bullshit *is* their evidence (I will return to this point later on).

Reisch argues that a pluralism about truth is foundational for bullshit (2006, 38). This, too, misses the mark slightly. A generosity about what counts as truth might aid a receptivity to bullshit, in cases where we are reluctant to reject a proposition out of hand despite not having any clear sense of what it means, especially if the bullshit in question trades heavily on skepticism about the possibility of knowledge, but it isn’t necessary: sophists have been around for a long time. What is dangerous here is not pluralism as such, but an audience’s abdication of the active project of understanding each other. Søren Kierkegaard’s pseudonymous works, for example, are replete with personalities who could be seen as engaged in several kinds of (often self-deceptive) bullshit.<sup>9</sup> But, what Kierkegaard is doing in presenting their viewpoints is emphatically not. The technique of indirect communication is a quasi-deception designed to bring the reader to empathise with and grasp the point of view of the pseudonymous

---

<sup>7</sup> See Jennifer Mather Saul, *Lying, Misleading, and What is Said: An Exploration in Philosophy of Language and in Ethics* (Oxford: Oxford University Press, 2012).

<sup>8</sup> Kimbrough also argues that “The most effective bullshitters know the truth, including the truth about when to bullshit and when to give the straight shit. The instrumental effectiveness of bullshit thus presupposes and exploits the instrumental effectiveness of truth...” Thus, effective bullshitting requires “a lively concern for the truth” (2006, 10). I take all bullshit to be instrumental; I would call what Kimbrough describes here as the eyes-wide-open variety, as opposed to either self-deceptive or desperate-evasive.

<sup>9</sup> See Davenport (2015).

personality and to see why they take certain propositions as true, in order to then respond *critically*—to see that they have gone down a (for Kierkegaard) wrong path, how they have done so, and for the reader to avoid it themselves. This exercise in perspective is foundational to a project of self-knowledge.

Bullshit is a conversational gambit, often launched so as to score (or avoid losing) social points. This can make the practice of bullshit resemble a game (Kimbrough 2006). To describe it this way, however, is to underplay what is at stake, both for the bullshitter and for the audience eventually inured to such tactics. Borman addresses this by way of the Habermasian distinction between strategic and communicative action. Bullshitters clearly fall into the category of strategic action, given that their utterances are calculated to procure certain social goods regardless of the cost to other social agents or to communicative social norms. Communicative action, by contrast, entails a commitment to “not only comprehending the utterances of our interlocutors, but also coming to a rationally motivated agreement regarding the content of those utterances” (Borman 2011, 122-3). Underlying this is an implicit agreement to participate in discourse as and presupposing that the other is a rational agent: one who operates on the basis of reasons and subject to rational justification. So, once again, we assume that our interlocutors aim at the truth, have justification for their claims, are largely rational, and mean what they say, because otherwise we cannot reasonably cooperate. The bullshitter violates this non-trivial norm of social interaction and does so in a way that parasitises it (2011, 125). Bullshit eventually undermines discourse as a means of reaching rational agreement. Thus, a consequence of the prevalence of bullshit and its public tolerance is the undermining of the “capacity of apparently competent speakers to make distinctions regarding the nature of language use which are integral to Habermas’ account of a communicatively integrated society”—which in turn undermines the possibility of making a rational critique of society (2011, 118).

### *Obfuscation*

Most of what has just been said about evasion can be said of obfuscation as well, which might be best understood in this context as a more specialised tactic of evasion. Nevertheless, it does have its own characteristics, ones that can be quite useful to the aspiring self-deceiver. As G.A. Cohen remarks, redirecting attention from speaker to product, “talking nonsense is what it is because of the character of its output, and nonsense is not nonsense because of features of the nonsense-talker’s mental state” (2002, 325). This is a fair objection to Frankfurt’s shift from talking about bullshit to talking about bullshitters. Cohen is right as well to point out that the falsity of bullshit as an artefact is at the level of tactic, whereas the “phoniness” of the bullshitter is at the level of goal (aim). The bullshitter attempts to convey something false, but the false thing that is said is not the specific falsity that is aimed at; indeed, the bullshitter may say true things in advancing this aim (2002, 330). Cohen’s particular concern, however, is less the bull than the shit, in this case, unclarifiable discourse “that is not only obscure but which cannot be rendered unobscure, where any apparent success in rendering it unobscure creates something that isn’t recognisable as a version of what was said”

(2002, 332).

Unlike Cohen, I am interested here in the bullshitter more than the bullshit, but the project of obfuscation is worth our attention here. In the academic cases Cohen considers, where theorists produce something that would be equally (im)plausible with or without an added negation (2002, 333), the next question should be “why so much obscurity?” Obviously, things often are just difficult and explanations tortuous because that is the nature of what is being explained. But where a speaker deliberately makes things much more obscure than necessary we can speculate as to possible motivations. Perhaps, the speaker/author doesn’t really understand the problem themselves but needs to come up with something that *looks* intelligent, or they think it will make them look more intelligent if no one can quite grasp what they are saying, with the further advantage that no one can prove them wrong—any objections can be dismissed with similar nonsense. They might think that what they have produced is particularly artful and that makes up for and justifies the obscurity, or that the obscurity demonstrates the ineffableness or fundamental indefensibility of the issues discussed, thus making a point about the impossibility of truth. Or, possibly, they think it really does make sense—indeed they must or else they have been wasting their career. All of these are familiar motivations for bullshit, and all of them, especially the last, also incentives for self-deception.

### ***Pretence***

Not all bullshitters are self-deceivers, nor need self-deceivers be bullshitters, but they frequently issue from similar motives, and bullshit offers techniques that are of significant practical value to self-deceivers. Both evasion and obfuscation serve as means of burying truths that the subject would prefer not to acknowledge and the self-deceiver needs some kind of cover for truths about or that affect that subject that the subject would rather were not the case, and which have some deep import for that subject. It is not that self-deception is always personal in the sense that it is about that person as such (it’s not about *me*), but we are far more inclined to motivational bias concerning those things in which we have some personal investment, whether emotional, intellectual, financial, political, cultural, and so on. Adherence to beliefs that are patently false about other people and things ought not to be instances of self-deception, except insofar as that adherence is integral to a strategy of blocking out realities of oneself. That is, while just being wrong about matters of fact doesn’t count as self-deception, there are ways in which factual error can become enmeshed in a self-deceptive project.

Normally, we deceive ourselves more or less directly about ourselves and our own situations, but a particularly complicated strategy of doing so could involve drawing in otherwise dubious beliefs about events, persons, or situations that, on the surface, have little if anything to do with us because maintaining those beliefs, and perhaps the amount of labour involved in supporting them, either justifies or distracts the lack of attention on matters closer to home. These beliefs serve to deflect attention away from the problematic area of one’s own life, character, behaviour, relationships, etc. If true, this would help us to understand why



some people seem impervious to clear evidence and compelling arguments against what seem to be irrational beliefs, as for example, with some of the more fantastical conspiracy theories, namely, that they hold those beliefs not only for the content of them but for what those beliefs do for *them*, and that this is why they seem self-deceived. If there is self-deception in such cases, it isn't because of *what* the subject holds or *why* they find *it*, specifically, to be epistemically attractive. Holding even grossly unfounded beliefs doesn't make a person self-deceived. What matters is whether the subject is simply misled or mistaken, or whether they have seized on a theory because it allows them to avoid consideration of something too crushing to accept. It is the deliberately engaging in self-anaesthetising bullshit and the fact that it does this work for the subject, rather than the particular variety of bullshit seized upon (e.g., homeopathy rather than conspiracy theory) that makes bullshit the handmaiden to self-deceit.

What makes scenarios such as these self-deceptive is the self-contrived avoidance of an issue about the subject and the use of external distractions to justify and facilitate that avoidance. In general, what is needed for successful self-deception is a way of keeping cognitive attention away from the distressing area—not so much that it ceases to have any effect on consciousness but enough to keep it from being explicitly acknowledged. The most obvious description of such a mental ploy is evasion. If all that is needed is evasion, the question of the actual truth or falsity of what is espoused and whether it is the negation of the truth from which attention is to be deflected is not irrelevant, but not critical. The mental filler just has to occupy the relevant “space” in order to provide a suitable diversionary substitute. Thus, if I need to avoid thinking about whether I am actually rather a jerk, I don't need to occupy my thoughts with why I am really quite considerate; getting exercised about all the ways that other people are jerks will do. But what that also implies is that a sufficient strategy for me is to engage in some opportunistic bullshit. Bullshit evades the question by occupying attention, throwing up explanations that can be just plausible enough or not at all, in order to distract from the truth or sincerity deficit of the speaker. It is verbal or cognitive stalling that, by travelling down a series of dead ends and irrelevancies, keeps the “audience” from concentrating on the question that matters, holding off self-examination.

It is here that Tamar Gendler's (2007) account of self-deception as pretence is most plausible. On Gendler's account, believing (or accepting) *p* as true is contingent on how one takes the world to be, but it also commits one to confirming one's belief in relation to the world, i.e., being willing to submit one's evidence for one's belief to rational scrutiny and abandoning it if there are grounds to believe it false (Gendler (2007), 236). Pretending that a proposition is true does not bear this evidential burden. In cases of self-deception, Gendler argues, the self-deceiver believes the true proposition but it does not guide their actions in ordinary circumstances, whereas, although they don't believe the false proposition, the pretence that it is true plays a belief-like role, i.e., by occupying thought and guiding one's interactions with reality (2007, 236-7). Thus, the self-deceiver acts *as if*  $\sim p$  is true.

The structure of self-deception Gendler lays out, then, is the following: **S** believes **p**; **S** does not believe  $\sim p$ , but **S** wishes that  $\sim p$  were true; **S** then pretends what things would be like if  $\sim p$  were true. A sufficiently sustained imaginative and performative practice of this kind of pretence might then, it is supposed, eventually result in belief (2007, 240-241). Gendler describes this situation as a substitution of a projective attitude where a receptive one is required. That is, although the subject has an abiding general commitment to holding truth-reflective thoughts or attitudes (thus a general commitment to being responsive to reality), they also have a specific reason to adopt or project a false or fictive one. The subject then needs to devise a way of “muting” the evidence for **p**, given their motivation for a  $\sim p$  world (2007, 242).

Pretending, however, is something that we do *as* pretending—it is a project that we deliberately set out to engage in. I cannot both pretend something and deceive myself about whether I am pretending. Moreover, if I gradually develop a belief in the pretence, it seems fair to describe this as *forgetting* what the truth is rather than pretending. To forget that the pretence *is* pretence is to no longer pretend but to accept as real.<sup>10</sup> Gendler’s analysis does point to the importance for self-deception of the way in which we explain reality to ourselves, particularly the way in which we leave out or deliberately alter certain aspects of that explanation; sometimes, indeed, pretending that they aren’t there. But pretence alone isn’t enough to make this work. If a self-deceiver does have both a basic commitment to reality and some motivation to not acknowledge some part of that reality, this may result in aversive behaviour but it will also require an explanation to self and possibly to others justifying the conflicting belief. This is a narrative project. If the self-deceiver does have a commitment to true belief, adoption of a course of action at odds with that commitment needs support in the form of a more or less worked out representation of what that person thinks the world is like and what they are doing in it. “I’d rather think of things this way” won’t do on its own; one needs to justify and incorporate one’s omissions and misrepresentations of evidence in a plausible and repeatable cognitive/linguistic account sufficient to convince both the self-deceiving subject and others, or sufficient to convince the subject that others are convinced. One needs to get the story straight, or straight enough, to be convincing. This story will have gaps, since parts are built over insufficiently evidenced pseudo-justifications, and eventually

---

<sup>10</sup> See also Gibson (2020, 674-5) on this. It is also not clear that Gendler’s examples (243-4) are really ones of self-deception rather than something like self-care or vanity (which admittedly is a motive for self-deceiving). I may avoid reminders of the way in which I am aging because they demoralise me, and focus instead on ways that I have retained youthful vigour, but that does not mean that I at any point think that I am not aging or that age isn’t having the effects that I find demoralising. Aversive behaviour engaged in because one does not want to be reminded of something that one knows, doesn’t amount to not believing it. It is more like a dreadful social event where everyone pretends not to hate each other for the sake of getting through it unscathed.

the gaps will become intolerably wide and require a new narrative.<sup>11</sup>

It is here that bullshit comes to the rescue of the committed self-deceiver, and it seems that what Gendler describes is less self-deception and more a sustained project of self-bullshit.<sup>12</sup> Bullshit is itself a kind of pretence, even when all the parties to its expression know that it is bullshit. Just as politeness has us nod in agreement with fictions for the sake of smooth social functioning, we let bullshit fly because we think it is not worth our while to call it out or because we find it amusing. In neither case do we believe that to which we assent. What the bullshitter is doing is performing, pretending to be more knowledgeable, or more generous, or more intimidating, etc. than they actually are. *Usually* this is a conscious action. There is, then, a certain overlap of bullshit with self-deception, in that bullshit works to maintain the security of the self-deceptive delusion although the content of that delusion is distinct from the truth-value phobic and deliberately obscuring character of active bullshit production.

### ***The Unreliable Self-Narrator***

Much of our sense of who we are is gained by our explanation of ourselves to ourselves and to others. This does not require that we have a complete and continuous self-history, nor does it require that it be the result of someone else's interrogation. Any self-accounting we do will be *ad hoc*, distorted by time and occasion, and constantly re-written in response to the pull of our current concerns. Nevertheless, this is what we do in order to make sense of where we are "now". If I want to understand why I want to do *this* with my life, why I feel *this* in these circumstances, why I always act in *these* ways, I endeavour to find an explanation that pulls together the threads of my own past and present, and accounts for *this* by connecting it with *that*. This may be an attempt at drawing out causal connections<sup>13</sup> or at giving rational justifications. This self-accounting will more than likely take on a narrative or quasi-narrative form. The "narrativity" here indicates that we tend to do this temporally or logically sequentially and more or less linguistically.

Self-narration, unfortunately, is inherently unreliable. As individual human beings we are biological organisms with definite material characteristics and histories, but as subjects we are also eminently capable of deceiving ourselves and, when we do, we are in precisely that

---

<sup>11</sup> See Fingarette (1969), 45-49. It might be reasonable to interpret Fingarette's description of the construction of an ever less plausible cover story as a type of bullshitting.

<sup>12</sup> This also raises the question of the extent to which engaging in a fiction and (self-) bullshitting overlap.

<sup>13</sup> Note that David Hume, while denying the existence of an abiding self, relies explicitly on these connections to explain the coherence of personality. *A Treatise of Human Nature*, (2007): 1.4.6, paragraphs 19-20, pp. 170-171.

condition where not only is certainty about ourselves lacking, but we are preventing ourselves from acquiring certainty. Someone who is self-deceived is the very person who cannot be trusted to be truthful about themselves. Even without self-deception, we frame our narratives of self according to context, audience, and the specific question to which we are responding. We don't offer exactly the same account of ourselves to our parents, our intimate partners, colleagues, teammates, casual acquaintances, and bosses. We tailor our accounts to fit what aspects of ourselves we judge suitable for show on that occasion, minimising or omitting some elements, introducing or exaggerating others. Consequently, our accounts are not strictly consistent.<sup>14</sup> We also speculate about our own motivations, often to fit what seems plausible to us now (supposing we are not also self-deceiving), and we are even more likely to do this where we simply do not know the truth about various aspects of our own histories.<sup>15</sup>

Part of what bullshit does in the service of self-deception is subvert the logic of self-accounting. If there is a way that we are, something that is true about us, then that should be accountable and explicable. But bullshit is all about making truth inaccessible. This is where obfuscation comes into its own as the best friend a self-deceiver could wish for. The point of self-directed obfuscating bullshit is to remove the truth standard from propositions about the self, so that (it seems) any proposition might just as well be true as any other, and therefore the one that we don't want to be true can't be, because truth can no longer be ascertained reliably, if at all, with respect to self. The self-deceiver can now assert themselves to be whatever and however they choose.<sup>16</sup>

The unclarifiable density of obfuscating bullshit makes this kind of self-deception more stable than simple evasion, such that it will be much more difficult for the self-deceiver to self-navigate out of their condition, and all the more so if many people are engaged in it. Borman's assessment seems particularly apt here: that in a society characterised by a prevalence of bullshit, individuals lose the sense that being in society comes with a normative dimension to

---

<sup>14</sup> Social media complicates this in some interesting ways, as what people state about themselves is visible across normal social boundaries of acquaintance. As a consequence, attempts to curate one's presentation to others as a single story seem fraught with difficulty.

<sup>15</sup> This is an epistemological rather than an ontological problem. I am not here suggesting that the self is in some sense ineffable, but only that our own *accounting* of who we are at any moment is potentially tricky. An important check on what stories we can reasonably tell about ourselves is provided by other people and by material reality as, whatever we take ourselves to be, we are *embodied* selves, always located in a specific biological organism, with a definite physical history. No self-narrative in which I am three metres tall or faster than Usain Bolt is going to meet that test.

<sup>16</sup> Fingarette observes that much of the problem here is about the identity we are willing to avow. See Fingarette (1969), Chapter 4, especially 65-74.

communicative interaction, “because its basis in the coordinating and integrating function of language oriented to reaching understanding has been swallowed up by language as spectacle, as strategy, as bullshit”—because the goal is not “rationally motivated conviction” but simply “‘getting someone to agree with you’” (Borman (2011), 133). Here, what the self-deceiver seeks is not self-knowledge or sincerity, i.e., well reasoned grounds for true belief, but affirmation of their projected interpretation of what they prefer to believe, or at least no sufficiently decisive disconfirmation. So long as the strategy is not defeated, it takes the place of rationally (and evidentially) motivated conviction.

If all this is correct, the situation is perhaps even more depressing than Frankfurt suggested. It is not simply that some people don’t care about the truth if they can get their way. Being cheated out of truth by the powerful provokes the outrage of epistemic injustice because we still believe in the possibility of truth. Self-deceiving bullshit is much more destructive—*unless*, that is, we are able to demand certainty that extends beyond the affirmation of personal projects of belief. In a Habermasian vein, this requires something closer to rationally motivated conviction on the basis of rational or empirical evidence mediated by a wide process of communicative interaction, which is itself governed by norms of rational and empirical justification. Without a strong and practical conviction in favour of such a view of rational communication and action, we really are lost.

But this also points to the end of the line for the inveterate self-deceiver—bullshit does eventually come up against material reality. Obfuscatory bullshit is a more stable strategy than the thinly evasive type for the self-deceiver, provided the goal is simply to avoid *acknowledgement* of the truth of the self-deceiver’s situation, but neither bullshit nor self-deception removes the truth that is thus disregarded. That is, you can bullshit all you like about how good your survival skills are but if you can’t read the map or start a fire when you are lost in the woods in winter, the bullshit won’t keep you warm or save your life. Because we have bodies, exist in a material universe that is independent of and indifferent to human narrative, and have factual and social histories that are intertwined with those of others, the project inevitably collapses in a pragmatic dead end; bullshit and self-deception are personally and socially destructive but ultimately futile attempts to rewrite who we are and who we want others to see us as. We do, in an important existential sense, make ourselves, but only in a context in which much of what we have to work with just isn’t up to us.

### **Conclusion**

I have tried to show that bullshit is not only a type of misleading speech that is commonly deployed by speakers in order to deflect attention or hide the truth from their audience. Most bullshit is a type of public utterance meant to misdirect attention away from one issue and towards another or, as in the case of performative sincerity, to persuade an audience that the speaker has certain normatively appropriate sentiments (or at least enough of them to realise that such a performance is required if not actually felt). While most discussion about bullshit has concentrated on the speaker’s relation to truth, I have argued

that we should pay more attention to what the speaker intends to accomplish. If we do, we find that bullshit has important strategic qualities that make it highly useful to self-deceivers and the wilfully self-oblivious. What bullshit can do in public discourse it can also do in internal narrative—it can evade and obfuscate truth. But whereas bullshit can, in principle, be called out in public discourse, though that may prove socially costly, the self-deceiver who bullshits inwardly has little in the way of internal check—and so a bullshitting self-deceiver who is also powerful is potentially a significant social danger. For the rest of us small fry self-deceivers, though, there is only the inevitable shock of discovery that neither we nor the world is quite as we thought.

## Works Cited

- Bernall, Sara. (2006). "Bullshit and Personality", in Gary L. Hardcastle and George A. Reisch (eds.) *Bullshit and Philosophy* (pp.63-82). Chicago and LaSalle: Open Court.
- Borman, David A. (2011). "Bullshit, Social Integration, and Political Legitimation: Habermasian Reflections". *Dialogue* 50, 117-140.
- Carson, Thomas L. (2009). "Lying, Deception, and Related Concepts", in Clancy Martin (ed.) *The Philosophy of Deception*, (pp. 153–87). Oxford: Oxford University Press.
- Cohen, G. A. (2002). "Deeper into Bullshit", in Sarah Buss and Lee Overton (eds.) *Contours of Agency: Essays on Themes from Harry Frankfurt*, (pp. 321-339). Cambridge, Mass: MIT Press.
- Davenport, John. (2015). "Frankfurt and Kierkegaard on Bullshit, Wantonness, and Aestheticism: A Phenomenology of Authenticity", in Anthony Rudd and John Davenport (eds.) *Love, Reason, and Will: Kierkegaard after Frankfurt* (pp. 73-113). London: Bloomsbury Academic.
- Fingarette, Herbert. (1969). *Self-Deception*. New York: Humanities Press.
- Frankfurt, Harry G. (1998). "On Bullshit". In *The Importance of What We Care About* (pp. 117-133). Cambridge: Cambridge University Press.
- Gendler, Tamar Szabó. (2007). "Self-Deception as Pretence", *Philosophical Perspectives*, 21, *Philosophy of Mind*, 231-258.
- Gibson, Quinn Hiroshi. (2020). "Self-Deception as Omission", *Philosophical Psychology*, 33, 5, 657-678.
- Hume, David. (2007). *A Treatise of Human Nature*, David Fate Norton and Mary J. Norton (eds). Oxford: Oxford University Press.
- Kimbrough, Scott. (2006). "On Letting it Slide". In Gary L. Hardcastle and George A. Reisch (eds.) *Bullshit and Philosophy* (pp.13-18). Chicago and LaSalle: Open Court.
- Lear, Jonathan. (2005). "Whatever", *The New Republic*.
- Mele, Alfred R. (2001). *Self-Deception Unmasked*. Princeton: Princeton University Press.
- Reisch, George A. (2006). "The Pragmatics of Bullshit, Intelligently Designed", in Gary L. Hardcastle and George A. Reisch (eds.) *Bullshit and Philosophy* . Chicago and LaSalle:

Open Court: 33-47.

Sartre, Jean-Paul. (1966). *Being and Nothingness*. Trans. Hazel E. Barnes. New York: Washington Square Press.

Saul, Jennifer Mather. (2012). *Lying, Misleading, and What is Said: An Exploration in Philosophy of Language and in Ethics*. Oxford: Oxford University Press.

Williams, Bernard. (1973). "Deciding to Believe," in *Problems of the Self*. Cambridge: Cambridge University Press.