

## On Kantian Bullshit

Immanuel Kant's derision of lying in his *Metaphysics of Morals* (hereafter *MM*) is among the harshest I have read on the topic. Kant spares no ink in vilifying what he categorizes as "the greatest violation of a human being's duty to himself regarded merely as a moral being" (Kant, *Metaphysics of Morals*, 6:429), robustly defending a duty to truthfulness in almost every instance. Harry Frankfurt, in relative congruence, opened philosophical inquiry into a tangential yet clearly distinct category of deception in 1986 with his paper (and subsequent 2005 book of the same name<sup>1</sup>) "On Bullshit." Bullshit, Frankfurt argues, is "a greater enemy of the truth than lies are" (Frankfurt, *On Bullshit*, p. 61) on account of its flippant disregard for truth in its entirety, in contrast with lying's predication on truth-relevance. Of 20th-Century origin, per Merriam-Webster Dictionary,<sup>2</sup> 'bullshit' was not a relevant deception category when Kant was developing his theses, so there is (expectedly) no mention of the practice in his theories of morality. Given his conviction for truth and Frankfurt's proclamation of bullshit's threat to truth, however, a critique of bullshit seems necessary for the Kantian moralist looking to extend Kant's reverence for truth-telling into the contemporary.

In the following essay, I delve into Frankfurt's bullshit with Kant as my guide, seeking to clarify the ways in which bullshit can manifest in light of Kantian work on lying, and in doing so look to provide the Kantian moralist tools with which to critique the practice. To do this, I begin with a relevant summation of Frankfurt's conception of bullshit and Kant's conception of lying, making particular note of how the Kantian definition of lying can be assimilated into the Frankfurtian conception of bullshit. I then look to develop a distinctly Kantian account of four manifestations of Frankfurt's bullshit, providing a taxonomy of bullshit while analyzing how

---

<sup>1</sup> All citations of Frankfurt's work of this title will be of the book *On Bullshit*, as opposed to the paper.

<sup>2</sup> Merriam-Webster. (n.d.). *Bullshit definition & meaning*. Merriam-Webster.  
<https://www.merriam-webster.com/dictionary/bullshit>

each subspecies should be viewed in conversation with Kant's moral framework. With a Kantian account of bullshit fully formed, I turn to evaluate Frankfurt's assertion that bullshit is a greater threat to the truth than lying, grounding my evaluation in the Kantian account I have developed. I then engage with G.A. Cohen's conception of bullshit, contrasting it with Frankfurt's account and evaluating Cohen's objection to the idea that bullshit can have a true truth-value, as well as Frankfurt's subsequent response. I conclude with a brief consideration of Frankfurt's suggestion that democracy necessarily produces propensity towards bullshit, and the potential considerations the Kantian might be wise to make in face of this.

## **1. Frankfurt and Kant: A Comparison of Deception**

### *1.1 Kant on Lying*

The first mention Kant makes of lying in the *MM* is when laying out the notion of 'freedom' as the original right to which every person has "by virtue of [their] humanity" (Kant, *MM*, 6:237). When it comes to external freedom (of which Kant's *Doctrine of Right* is primarily concerned with), utterances of intentional falsehood rise to the level of a lie only if they "directly [infringe] upon another's right" (Kant, *MM*, 6:237. fn. 26). An instance of this, for example, would be the telling of untruths to someone about the content of a legal document you then have them sign. For Kant, such an action directly denies one's ability to be free from others' choices. This is notably weak when compared to the opening quotation from his *Doctrine of Virtue*, however it aligns quite well with the difference in domains that Kant views right and virtue to have. Since right is concerned with *external* freedom—i.e. those freedoms which should be dictated and protected by civil law—the severity of the lie must rise to a level which the law

should be concerned with. Virtue's concern with *internal* freedom, on the other hand, allows it to be much more fine-grained with its condemnation of lying.

While lies that deny external freedom such as libel or slander fall under the jurisdiction of the *Doctrine of Right*, the intentional untruths that do not rise to the level concerning the law find their moral contemptuousness defined in the *Doctrine of Virtue*. For Kant, what distinguishes lying as morally reprehensible cannot merely be the harm it causes others (this is a concern of the *Doctrine of Right*) nor the harm it causes the liar's self (this is a concern of the pragmatic maxim). Instead, lying's moral transgressions are a matter of internal freedom—a rejection of the self as human. “A human being who does not [themselves] believe what [they tell] another...has even less worth than if [they] were a mere thing” (Kant, *MM*, 6:429). Kant's moral philosophy is predicated on the notion that each and every person is an end unto themselves, and should be treated as such—not as mere means by others. An act of lying defiles this notion. Instead of following the natural obligation one has to truthfulness—to which one's inner end of communicating thoughts is tethered—lying renders *oneself* as mere means in the form of “a speaking machine” (Kant, *MM*, 6:430). The conscience,<sup>3</sup> when taken with the utmost seriousness as Kant implores, is akin to another person whose violation is just as wrong. An inner lie, then, is one in which one's conscience is treated as a means rather than an end to itself, violating the duty one has to themselves as a moral being.

Elsewhere, in his response to objections from Benjamin Constant titled “On a Supposed Right to Tell Lies From Benevolent Motives” (hereafter “Supposed Right to Lie”)<sup>4</sup> Kant deepens his commitment to truth by grounding it as a duty to others. When one lies to another, “all rights founded on contract...lose their force” (Kant, “Supposed Right to Lie”, p. 1). Further on, Kant

---

<sup>3</sup> “inner judge” (Kant, *MM*, 6:431)

<sup>4</sup> The title of this piece has also been translated as “On a Supposed Right to Lie from Philanthropy”.

reinforces this by asserting a “strictest duty to truthfulness in statements which [one] cannot avoid” (Kant, “Supposed Right to Lie”, p. 2) even if said true statements have the capacity to do harm or danger to the speaker or others. Given that all of this is happening in the context of telling the truth to a murderer about their target’s whereabouts, it is important to note that Kant is not claiming that telling a lie in this instance wrongs the murderer. Helga Varden has defended this notion in her work “Kant and Lying to the Murderer at the Door...One More Time: Kant’s Legal Philosophy and Lies to Murderers and Nazis” (hereafter “Kant and Lying”) where she parses out the difference in Kant’s analyses of right (justice) and virtue (ethics). Here, Varden notes that “a person simply does not have a right against another person that he tells her the truth” under Kant’s theory of right laid out in the *MM*. Recall that under external freedom, an intentional untruth only rises to the level of a lie if it infringes on another’s freedom. Any utterance that is intentionally untrue does not necessarily rise to this level as, as Varden puts it, “words in general do not have coercive power on Kant’s view” (Varden, “Kant and Lying”, Sec. 3). The notion that the liar in the murderer case has done wrong is then not that they have wronged the murderer, but that they have wronged *humanity* by upending the formal principle of truthfulness we have to “a world of rightful interaction” (Varden, “Kant and Lying”, Sec. 3). For the purposes of this paper, this notion of wrongdoing qua virtue on the part of the liar is the salient one, as opposed to the stronger conditions needed to classify an intentional untruth as a lie qua right.

Notably, Kant denies that a lie is present when no one is deceived by the content of the lie (including the speaker). For instance, he gives the case of an author who asks a reader of their book how they enjoyed it. The reader, knowing any hesitation on their part will be viewed negatively by the author, voices their approval as that is what is expected of them in the situation.

Despite the possibility of this utterance being untruthful, however, there can be no deception, as there is no expectation that what the reader responds with holds any genuine truth value. The author and reader here are much like the American who asks the cashier “How are you today?” The expectation is a positive response, not the truth, so no lie (which is predicated on deception) can be told. More will be said on this in the discussion on bullshit.

### *1.2 Frankfurt on Bullshit*

To illustrate the type of speech that he is attempting to classify in *On Bullshit*, Frankfurt grounds the discussion in a purported conversation between Fania Pascal and Ludwig Wittgenstein from their time at Cambridge in the 1930s:

“I had my tonsils out and was in the Evelyn Nursing Home feeling sorry for myself. Wittgenstein called. I croaked: “I feel just like a dog that has been run over.” He was disgusted: “You don’t know what a dog that has been run over feels like.” (Frankfurt, *On Bullshit*, p. 24)

While this seems like at best an attempt at deprecating humor by Wittgenstein, and at worst a unreasonable complaint about a simple simile, it does quite a nice job of showing the type of speech that Frankfurt is getting at when taking the point of view of Wittgenstein (assuming him to be serious in his response). Wittgenstein’s primary objection is that Pascal is uttering a statement that is purported to convey information about how she feels—an act that is reliant on truth/false distinctions—however Pascal seems to have no interest in whether what she has actually said is true or false. This lack of concern for the truth-value of her utterance is what Frankfurt describes as “the essence of bullshit.”<sup>5</sup> Pascal does not know what it feels like to be a

---

<sup>5</sup> It should be noted that Frankfurt partially walks back this claim of “essence” in his response to G.A. Cohen’s “Further Into Bullshit.” In the response, Frankfurt notes that his essence claim is misleading in its suggestion that all genuine instances of bullshit must contain this intentional disregard for truth-value, though he does not recant the “essence” description. Since this paper is primarily focused on Frankfurt bullshit—of which this essence is seemingly necessary—I have kept the reference as such.

dog that has been run over, and so she has no ability to ascertain the truthfulness of the comparison of her post-tonsil removal state to that of the ran-over dog. Nonetheless, she compares the two without a care for the truth-value, as her only motivation is to effectively convey that she feels poorly.

This disregard for truth-value within an utterance is the crux of Frankfurt's separation of bullshit and lying. When it comes to the liar, they *must* be directly concerned with truth, and indeed must know what is true in order to successfully lie. If the liar has no conception of truth, they are unable to "design [their] falsehood under the guidance of that truth" (Frankfurt, *On Bullshit*, p. 52). In a sense, the liar has a reverence for truth and places importance on it, positioning themselves directly against that truth. The bullshitter, in contrast, "pays no attention to it at all" (Frankfurt, *On Bullshit*, p. 61). The truth or falsity of their speech is inconsequential, as the bullshitter is singularly concerned with a pursuit of their favorable ends.

This leads to a quite interesting wrinkle, which is that the bullshitter may indeed be telling the truth. A helpful comparison Frankfurt provides on this notion is that of the con artist to the bullshitter. The issue with a con artist is not that the fake work is *bad*—in some cases the works are incredible to the point that they're indistinguishable from the original—but that it is *fake* (Frankfurt, *On Bullshit*, p. 47). Similarly, the bullshitter does not necessarily deceive their victim on grounds of falsity, but of fakery. For Frankfurt, the bullshitter "may not deceive us, or even intend to do so, either about the facts or about what [they take] the facts to be. What [they do] necessarily attempt to deceive us about is [their] enterprise" (Frankfurt, *On Bullshit*, p. 53–4). Since the truth-value of the bullshitter's statement is wholly inconsequential to them, the wrong that they are committing cannot be grounded in misrepresentation of truth-value, but deception of aims. The bullshit that is spewed might very well be true, similar to the way in

which the counterfeit art piece might very well be high quality art. Nonetheless, the wrong that is done is the misrepresentation of enterprise on the part of the bullshitter, as it is for the con artist. The bullshitter conceals their true aim under the bullshit, the same way the con artist conceals the true origin of their artwork under the artwork itself.

Because the bullshitter is coercive in this fakery sense, and has more creative means at their disposal with which to deceive when compared to the liar, Frankfurt views bullshit as a greater threat to truth than lying. The liar must have a sense of reverence for the truth, as they are “inescapably concerned with truth values” (Frankfurt, *On Bullshit*, p. 51). In order to craft a successful lie, the liar necessarily must know the truth, or else they cannot position themselves against it. The bullshitter, on the other hand, has no regard for truth-value, and wantonly casts it aside in pursuit of their ends. They “[do] not care whether the thing [they say describes] reality correctly. [They] just pick them out, or make them up, to suit [their] purpose” (Frankfurt, *On Bullshit*, p. 56). In this sense, bullshit is indeed more insidious than lying, as lying at least has some respect for the truth and grounds itself upon this respect.

### *1.3 Kantian Lying within Frankfurtian Bullshit*

With accounts of lying and bullshit provided by Kant and Frankfurt, respectively, we can now turn our attention to the question of whether Kant’s version of lying is compatible with the way Frankfurt describes lying in *On Bullshit*. If the two constructions of lying are incompatible, then it seems we cannot adequately evaluate Frankfurtian bullshit on Kant’s view without radically revising the ways in which Frankfurt compares lying and bullshit throughout his work. After all, his juxtaposition of lying and bullshit with respect to truth is predicated on his

definition of lying. If Kant's approach to lying cannot be similarly juxtaposed due to fundamental differences in conception of lying, then we have much more work to do.

Luckily for us (or me, particularly), on lying, Kant and Frankfurt seem to develop quite similar accounts, with both involving intentional misrepresentation of the truth, positioning the liar as diametrically opposed to truthfulness. It should also be noted that Frankfurt has given credence to Kant's notion of lying in his essay "On Truth, Lies, and Bullshit," however he makes a point of saying that the gravity that Kant assigns lying with respect to humanity's cohesion is "exaggerated" (Frankfurt, "On Truth, Lies, and Bullshit, pt. II). Nonetheless, this does give us solid basis to assimilate Kant's framework of lying into Frankfurt's conception of bullshit, given that the way in which lying operates seems to be endorsed as compatible for Frankfurt (while it is the intensity of the harm it causes which he objects). Given that I am mostly interested in how the Kantian views bullshit, as opposed to how the Frankfurtian views lying, this compatibility is enough to allow us to get the project off the ground. It will be important to note this difference in the perceived gravity of lies, however, when evaluating to what extent (if any) bullshit is worse than lying for Kant, as we cannot just assume that he would agree with this assertion of Frankfurt's given their noted differences on lying.

## **2. On Kantian Bullshit**

I propose the following taxonomy of Frankfurtian bullshit in order to categorically evaluate what I believe to be four subspecies of bullshit as Frankfurt conceives, predicated on what the relative truth-value of the bullshit is, and whether or not the bullshitter knows the truth-value of the bullshit when uttered:



	VALUE KNOWN	VALUE UNKNOWN
TRUE	(I) Bullshit is true and the speaker knows it	(II) Bullshit is true and the speaker does <i>not</i> know it
FALSE	(III) Bullshit is false and the speaker knows it	(IV) Bullshit is false and the speaker does <i>not</i> know it

Table A. Taxonomy of Bullshit

My motivation for introducing this taxonomy is twofold. First, I believe that this categorization will allow for a clean evaluation of bullshit through a Kantian lens, given the importance Kant places on truthfulness in his works. By organizing subspecies of bullshit by what the truth-value of the bullshit is and whether or not the speaker of the bullshit knows the truth-value, we are able to more clearly ascertain how Kant might have conceptualized bullshit with respect to duty. In other words, conceptualizing bullshit in terms of truth-value gives us access to a greater range of tools in our Kantian toolkit. The second motivation is that—given the deep importance Kant places on truth and intention when discussing lying—it seems that the respective truth-value of the bullshit, as well as whether the speaker knows of this truth value will necessarily factor into how Kant would condemn the practice. Whether the bullshit is true is certainly relevant, as that will determine whether or not the speech has the capacity to create a false reality for the victim of the bullshit. Similarly, whether or not the bullshitter knows the truth-value of the bullshit seems highly relevant to whether a greater degree of wrongdoing can be assigned to them. This taxonomy allows the Kantian to evaluate each instance of bullshit independently, allowing for degrees of wrongdoing to be assigned depending on the severity. This is in line with Kant’s conception of evil, which he defines as “deviation of the maxims from

the moral law” and divides into three distinct degrees—frailty, impurity, and wickedness (Kant, “Religion within the Limits of Reason Alone”, Bk. 1, Ch. II). By denoting four subspecies of bullshit, we are able to evaluate them more finely, as Kant does lying.<sup>6</sup>

### *2.1 A Kantian Evaluation of Bullshit Subspecies*

With a taxonomy established, we are now at a place to evaluate how Kant might have conceptualized bullshit had it been a prominent deception category during his time. Before jumping in, however, I want to briefly discuss a formulation of bullshit Frankfurt describes in which the Kantian might want to resist claiming there is a violation of duty. Recall the anecdote told by Fania Pascal about her phone call with Ludwig Wittgenstein. In this case, it seems clear that Pascal did not intend deception, despite Wittgenstein’s seeming response to the contrary. I believe the Kantian is best served by putting this sort of figurative form of bullshit into the same conceptual bucket as the reader who commends the author on their book. Nikil Mukerji and Adriano Mannino take a similar approach in their discussion of argumentative bullshit, stating in a footnote that this example of “lying” from Kant is tantamount to a tacit endorsement of bullshit (Mukerji & Mannino, “Deeper into Argumentative Bullshit”, p. 441, fn. 2). While I agree in this specific instance of Pascal and Wittgenstein that the Kantian should see bullshitting of this form as acceptable, I vehemently disagree with their sweeping generalizations of all bullshit. As we shall see, there are plenty of forms of bullshit the Kantian can—and should—condemn.

---

<sup>6</sup> Kant, for instance, defines one’s wish to see only the good in their partner, which blinds them from seeing faults, as an example of an inner lie that amounts to frailty, but nothing more (Kant, *MM*, 6:431).

### 2.1.1 (I) Bullshit is True and the Speaker Knows It

Bullshit of Type I is fascinating in that it seems to conceptually contradict itself. How can one be bullshitting if what they are saying is true and they know it to be true? Similarly, when evaluating from a Kantian perspective of lying, it initially seems to escape unscathed. After all, the issue Kant has with lying is that it is an intentional untruth, however there is nothing untrue about Type I bullshit—it is, by all accounts, true.

To help with the conception of what Type I bullshit might entail, consider the undergraduate (or, in this case, the graduate<sup>7</sup>) writing a paper about Kant. It is certainly possible to see the student write dense, meaningless dribble that repeatedly and vaguely invokes things like the Categorical Imperative and Kant's propensities towards evil, without actually saying anything at all. The paper as a whole might be utterly nonsensical, but not *false*. After all, the information and the quotes used in it seem true, and on the whole the student seems to understand this. Nonetheless, the student isn't concerned with the truth value of the paper at all. Their theorization could be completely false for all they care, they just see the paper as means to a good grade, and their bullshit is predicated on attaining this and nothing else. This type of bullshit is what I believe can be classified as Type I.<sup>8</sup>

With regards to Kant, it seems clear that we cannot evaluate this in congruence with his condemnation of lying, as there are no untruths involved. Nonetheless, it does seem as though this sort of bullshit does rise to a violation of the maxim of the will, as there is some sense of disregard for the professor for whom the paper is submitted to, as well as the writer themselves,

---

<sup>7</sup> While I certainly hope that this paper is not a load of bullshit and, in fact, actually has something meaningful to contribute to both the bullshit and Kantian canons, there is some aspect of writing on bullshit that feels, well, bullshit. Nevertheless, I persist.

<sup>8</sup> This and similar instances of bullshit in academia are of particular concern to G.A. Cohen in his response to Frankfurt: "Further Into Bullshit."

as ends. The best place to find basis for this, then, would be in one of the other vices contrary to moral duty. I suggest ‘servility’ as a possible home for this type of maxim violation.

In discussion of servility, Kant contends that “belittling one’s own moral worth merely as a means to acquiring the favor of another, whoever it may be...is false (lying) humility, which is contrary to one’s duty to oneself since it degrades one’s personality” (Kant, *MM*, 6:436). This seems in line with the student who, using themselves as mere means to a good grade, belittles their own capacity for knowledge production and truth-evaluation in the process. This, in turn, degrades their personality to that of a bullshitter, leading the student to be guilty of “morally-false servility,” which violates the duty of humility one has to oneself as a moral being. This deviation from virtue most likely results in frailty when categorized using Kant’s degrees of propensity to evil. If the student is a serial bullshitter in papers, however, it could certainly rise to impurity and possibly wickedness, though I am, at this moment, skeptical of the latter.

### *2.1.2 (II) Bullshit is True and the Speaker Does Not Know It*

Along with Type IV, Type II bullshit seems to be the most common of the bullshit that floats through the world. Consider, for instance, the politician who makes a (correct) assertion of a drastic increase of crime based on anecdotal evidence that they lay out in a speech. Telling a story of a woman who was robbed while out jogging, the politician declares that the city has a crime surge, calling on voters to vote them into office in order to increase the number of police patrols per night. This politician is wholly unaware of the truth-value of their statement—they simply made the claim because they knew it would help their election bid. Nonetheless, the statement is true. Similar to Type II bullshit, this cannot be a violation in the form of lying, as

there are no untruths present. So here, we again turn to Kant's discussion of servility to evaluate potential violations of duty.

As with Type I bullshit, I believe that the Kantian is best served by condemning this form of bullshit through classifying the practice as a degradation of humility. The politician, in casting aside truth value in their speech in search of votes, is certainly belittling their moral worth in an attempt to gain the favor of others. Kant says to "be no man's lackey" (Kant, *MM*, 6:436), yet the politician who casts aside the truth to curry votes in their favor lowers themselves to the place of a lackey, and in doing so violates their duty to themselves as a moral being. This subspecies of bullshit is much more likely to cross the threshold into impurity, as it becomes a reliable mechanism to garner votes in exchange for selling one's humility. Wickedness results when the politician convinces themselves that this is for their own benefit, which then "neglect[s] the incentives springing from the moral law in favor of others which are not moral" (Kant, "Limits of Reason Alone", Bk. 1, Ch. II) —that immoral law being servility.

### *2.1.3 (III) Bullshit is False and the Speaker Knows It*

Type III bullshit is the clearest account for which the Kantian has the requisite tools to condemn, on account of its function as a lie as well as bullshit. Consider the case of an advertiser who chooses to make a false claim in their advertisements, but are motivated to do this by the effect the advertisements will have on revenue, as opposed to a goal to mislead consumers. Frankfurt contends that "in that case what they choose to convey is something that they know to be false...they end up not merely bullshitting but telling lies as well" (Frankfurt, "Reply to G.A. Cohen", p. 341). While the goal of the bullshitter and the liar is inherently different—one is directly concerned with truth and one is wholly dis-concerned with it—a case of Type III bullshit

still results in the bullshitter being guilty of telling lies as well. Given that there is an intentional untruth—the Type III bullshitter is intentionally making a statement they know to be false, even if they are unconcerned with the truth value as a means to their ends—this subspecies of bullshit is firmly derided by Kant as a violation of duty to oneself as a moral being on account of defiling truthfulness. If anything, there may be added consideration of whether or not the Kantian should view Type III bullshit as *more* egregious than a simple lie, as they refuse to treat truth with the reverence it deserves, *and* directly opposes it in action. In this sense, it might be considered a violation of lying and of servility, if we were to employ a similar argument as in types (I) and (II). With regards to propensity of evil, it seems that Type III bullshit has the capacity to occupy all three degrees, as lying does for Kant.

#### *2.1.4 (IV) Bullshit is False and the Speaker Does Not Know It*

As stated in discussion of Type II, I consider Type IV bullshit to be among the most prolific of the subspecies taxonomized in this paper. Frankfurt asserts that “bullshit is unavoidable whenever circumstances require someone to talk without knowing what [they are] talking about” (Frankfurt, *On Bullshit*, p. 63). I agree, and would go as far to say that when these circumstances arise, the odds that the relative truth value of the bullshit is false is much greater than it being true. I say this on account of the way in which we tend to consider truth and falsity in classical logic, where to declare something false, one simply needs it to be false in *one* set of circumstances. Conversely, for something to be true, it must be true in *all* circumstances. Because of this, declaration of truth is much harder to attain than declaration of falsehood. I see no reason for it to be different when bullshit is the object of inquiry, though I leave open the

potential for it to be the case, as it has little to no bearing on how the Kantian should condemn this sort of bullshit.

To evaluate a Type IV bullshit utterance, consider the following statement made by then-presidential candidate Donald Trump during a debate with fellow candidate Kamala Harris on September 10th, 2024: “When you look at these millions and millions of people that are pouring into our country monthly...and just look at what they’re doing to our country. They’re criminals, many of these people are criminals, and that’s bad for our economy too.”<sup>9</sup> The claim that “millions and millions of people” are entering the United States monthly is verifiably false, as the Associated Press’ fact-checking wing has shown. They also contend that it is unproven that “many of these people are criminals,” noting that many studies have found no link between immigration and increased crime. The sentence was uttered on the debate stage in order to convince those who view increased immigration as a bad thing to vote for Trump, with no consideration for the truth-value of the statement whatsoever. If we assume Trump was not aware of the truth-value of the sentences, then it seems that the above quotation is a clear-cut case of Type IV bullshit.

When evaluating this subspecies of bullshit under the Kantian framework, it might best serve us to begin with lying. This type of bullshit is the intentional utterance of what can be classified qua truth-value as an untrue statement, but does it classify as an “intentional untruth” in the way Kant conceives of it? While Kant does not explicitly define what he means by the term “intentional untruth” to my knowledge, he gives a decisive clue to how the phrase functions in the “Supposed Right to Lie”, where Kant defines a lie as “an intentionally false declaration towards another” (Kant, “Supposed Right to Lie”, p. 1). This usage of ‘intentionally’ as an

---

<sup>9</sup> Associated Press. (2024, September 11). Fact focus: A look at false and misleading claims made during Trump and Harris’ debate. AP News.  
<https://apnews.com/article/debate-harris-trump-fact-focus-abortion-crime-0e368102d38d87d708a8a497a38f4fb2>

adverb instead of the adjective usage previously mentioned implies a direct relation to the falsity of the statement. Here, ‘intentionally’ is acting directly on ‘false’, requiring that the statement be not only intentional qua statement, but intentional qua *false statement*. Extrapolating this to his usage of ‘intentional untruth’ in the *MM*, we can now view the statement as *intentionally untrue*.

This valuation seems to let the Type IV bullshitter off the hook when it comes to a critique on the basis of Kantian lying. After all, they are unaware of the truth-value of their statement, and therefore cannot be intentionally stating falsehoods. In fact, the bullshitter would rightly deny that they are being intentionally false, as they are wholly unconcerned with truth-value, and could not care less whether their bullshit is true or false. However, it is this flippant disregard for truth-value in combination with a false bullshit statement that allows the Kantian to reel them back in.

Later on in the “Supposed Right to Lie”, Kant asserts that “to be truthful (honest) in all declarations is therefore a sacred unconditional command of reason, and not to be limited by any expediency” (Kant, “Supposed Right to Lie”, p. 2). This command of reason sinks the Type IV bullshitter, whose disregard for truth value—and truth as a result—flies in the face of this sacred and unconditional command. The individual who believes a lie told to them, and then purports it later, is able to skate around using this claim, as reason dictates that they were right to believe the lie, even if it was indeed false. The bullshitter has no such trump card, as they never cared for the truth value of their statements in the first place. By casting truth to the side, the Type IV bullshitter loses their defense of ignorance, as they have already denied the command to be truthful in all declarations. Trump is rightfully condemned by the Kantian for his Type IV bullshit—even if the falsity of his statements is not known to him—because he has already transgressed reason by disregarding truth from the outset.



### 3. Conclusion

#### 3.1 *On The Gravity of Bullshit and Lying*

If you recall, at the beginning of this paper I provided two quotes from Immanuel Kant and Harry Frankfurt, respectively. The former derided lying as the greatest threat to truth, while the latter asserted that bullshit is truth's greatest enemy. Frankfurt was able to make that declaration with a notion of lying and bullshit in mind, while Kant was not. With this Kantian evaluation of bullshit, however, we are now seemingly in a position to at least consider the possibility that bullshit is a greater threat to Kant's truth than the lying he so deeply disdained.

Given the four subspecies of bullshit analyzed, I am skeptical of Frankfurt's claim if applied to a Kantian conception of truth and lying—at least for three of the four subspecies. Types I and II seem to not be a threat to truth much at all, given that they are indeed true. There is an argument to be made that they degrade truth by relegating it to a lesser status within the communicative exchange, but I don't see that as being a greater threat than the undermining of truth that lying operates on. Type III is bullshit in cahoots with lying, and so provides a dual threat to truth, but it seems difficult to say that in this instance bullshit is doing the heavy lifting when it comes to degradation of truth. An argument could certainly be made, but at this point I am agnostic to the claim of one over the other. While truth is also relegated in this instance, it is the lie that directly takes it on, and so the gravity of the relegation over direct denial would need to be made salient in order for me to get on board.

A Type IV bullshit statement is where I believe Frankfurt has some serious ground to stand on. Like its fellow subspecies, Type IV bullshit relegates truth to a lesser status in communicative exchange. This relegation is more pernicious, however, due to the nature of Type IV bullshit as unknowingly false. The bullshitter can proliferate falsehoods while genuinely

denying that what they are saying is false at all, and indeed even scorning the notion that they are lying. After all, they have no knowledge of the truth-value of their statement, and have no interest in it. This allows them to lock truth in the communicative basement while parading as its champion, effectively countering calls that they are a liar while spewing false bullshit with wanton abandon. In this instance, it seems wholly appropriate to declare bullshit a greater threat to truth than lying, for at least the liar shows respect to truth before opposing it. Type IV bullshit is guerilla warfare while lying is proper war etiquette; is one more threatening than the other? I don't know, but I certainly would prefer to face my opponents head-on than deduce what nook or cranny they are hidden in.

### 3.2 G.A. Cohen and "*Deeper Into Bullshit*"

G.A. Cohen penned a notable response to Frankfurt in 2002 that I would be remiss not to discuss, particularly given Frankfurt's subsequent response to Cohen contains a key defense of two of my subspecies of bullshit. In "*Deeper Into Bullshit*", Cohen makes a distinction between "output-centered" bullshit and "activity-centered" bullshit, the latter of which he takes Frankfurt's account to be dealing with (Cohen, "*Deeper Into Bullshit*", p. 331). More particularly, Cohen believes that the distinction between bullshitter and bullshit is important, and takes his output-centered approach to draw out the distinction between the two in a way Frankfurt's neglects to do. The most notable claim Cohen makes with respect to this paper is that he denies that Type I and II bullshit is bullshit at all. On this, he states that "a person who speaks with Frankfurtian indifference to the truth might do so yet *happen* to say something true, and, in at least one sense of the term, the one that interests me, what he says could not then be bullshit" (Cohen, "*Deeper Into Bullshit*", p. 331, *italics original*). Given Cohen's interest in solely the

output, a true statement cannot be bullshit, as it is true. He has a similar paradigm case for a person who utters bullshit, yet is not a bullshitter.

Frankfurt defends his account in his response, stating that, for him, “the mental state of the person who creates some piece of discourse is a crucial factor in determining whether or not what is created is bullshit” (Frankfurt, “Reply To G.A. Cohen”, p. 341). I am inclined to agree with Frankfurt, and reject the notion that the simple fact that what someone is saying is true is able to save it from being labeled bullshit without further explanation as to why that would be the case. As in the two examples I gave for Types I and II, it certainly seems enough for the truth-values to be disregarded to declare something bullshit, even if the statement itself is conveying truth. The bullshitter’s conveyance of truth is secondary to their manipulation of the communication exchange, and that is enough for me to declare the statement bullshit, irrespective of relative truth-value. Who can and should be given the label of bullshitter is an interesting question, but one that bears little weight on the discussion in this paper.

### *3.2 Bullshit and Democracy*

As a final thought to leave you with, Frankfurt makes note of the relationship between the duties citizens have to their democratic governments and bullshit. Recall his assertion that bullshit is inevitable when people discuss matters of which they are ill-informed. Frankfurt notes that “it is the responsibility of a citizen in a democracy to have opinions about everything, or at least everything that pertains to the conduct of [their] country’s affairs” (Frankfurt, *On Bullshit*, p. 64). Though I did not address it in this paper, the Kantian would do well to make note of how this might work into Kant’s political theory, particularly given his affinity for Enlightenment thought and the social contract tradition. If bullshit is an inevitable consequence of democracy, it

might be wise to devise ways to curb its influence as much as possible, particularly if the kind that festers is of Type IV.

### Bibliography

Cohen, G.A. “Deeper Into Bullshit”. Buss, S. (2011). *The contours of agency: Essays on themes from Harry Frankfurt*. The MIT Press.

Frankfurt, H. G. (2005). *On bullshit*. Princeton University Press.

Frankfurt, H. G. “On Truth, Lies, and Bullshit”. Martin, C. W. (2009). *The Philosophy of Deception*. Oxford University Press, pp. 37–48

Frankfurt, H. G. “Reply To G.A. Cohen”. Buss, S. (2011). *The contours of agency: Essays on themes from Harry Frankfurt*. The MIT Press.

Kant, I., Denis, L., & Gregor, M. J. (2017). *The metaphysics of morals*. Cambridge University Press.

Kant, I. (n.d.). On a supposed right to tell lies from benevolent motives. (1797)  
<http://philosophical.space/f325/KantLies.pdf>

Kant, I. (1996, January 25). *Religion within the limits of reason alone*.  
<https://staffweb.hkbu.edu.hk/ppp/rbbr/toc.html>

Mannino, A., & Mukerji, N., “Deeper into Argumentative Bullshit”. *Informal Logic* 42, no. 2 (2022), pp. 439–470.

Merriam-Webster. (n.d.). *Bullshit definition & meaning*. Merriam-Webster.  
<https://www.merriam-webster.com/dictionary/bullshit>

Associated Press. (2024, September 11). *Fact focus: A look at false and misleading claims made during Trump and Harris' debate*. AP News.

<https://apnews.com/article/debate-harris-trump-fact-focus-abortion-crime-0e368102d38d87d708a8a497a38f4fb2>

Varden, H. "Kant and Lying to the Murderer at the Door... One More Time: Kant's Legal Philosophy and Lies to Murderers and Nazis." *Journal of Social Philosophy* 41, no. 4 (2010), pp. 403–421.