

# Deeper into Bullshit

G.A. Cohen

\* \* \*

## 2. Two Species of Bullshit

I should like to explain how this chapter reached its present state. I read Frankfurt's article [i.e. Harry Frankfurt, "On Bullshit"] in 1986, when it first appeared. I loved it, but I didn't think critically about it.

Having been asked to contribute to the present volume, I reread the article, in order to write about it. I came to realize that its proposal about the "essence" of bullshit worked quite badly for the bullshit ... that has occupied me. So I wrote a first draft which trained counter-examples drawn from the domain of the bullshit that interests me against Frankfurt's account. But I then realized that it was inappropriate to train those examples against Frankfurt, that he and I are, in fact, interested in different bullshits, and, therefore, in different *explicanda*. Frankfurt is interested in a bullshit of ordinary life,<sup>1</sup> whereas I am interested in a bullshit that appears in academic works, and, so I have discovered, the word "bullshit" characteristically denotes *structurally* different things that correspond to those different interests. Finally, and, belatedly, I considered, with some care, the *Oxford English Dictionary* (OED) account of "bullshit", and, to my surprise, I discovered (and this was, of course, reassuring) that something like the distinct *explicanda* that I had come to distinguish are listed there under two distinct entries.<sup>2</sup>

So, instead of citing cases of the bullshit that interests me in disconfirmation of Frankfurt's account, I now regard it as bullshit of a different kind.<sup>3</sup> Which is not to say that I have no criticism of Frankfurt's treatment of the kind of bullshit that interests him.

---

<sup>1</sup> His essay begins as follows: "One of the most salient features of our culture is that there is so much bullshit. Everyone knows this. Each of us contributes his share" (p. 1).

<sup>2</sup> Frankfurt himself cites the OED, but mainly with respect to meanings and uses of the word 'bull': he touches on its definition of "bullshit" only in its use as a verb. I disagree with his discussion of the entries he cites, but it would be an imposition on the reader's capacity to endure tedium to explain why.

<sup>3</sup> Four differences between the kinds of bullshit that exercise Frankfurt and me are listed in footnote 16 below. The import of those differences will emerge in due course, but the reader will probably follow me better if he or she glances ahead now to footnote 16.

... [Frankfurt] begins by saying that the term 'bullshit' is very hard to handle, analytically, but, as we shall see, he rather abandons caution when he comes to offer his own account of it.

Consider, then, the OED reading of 'bullshit':

bullshit *n.* & *v.* *coarse sl. - n.* 1 (Often as *int.*) nonsense, rubbish. 2 trivial or insincere talk or writing.<sup>4</sup> - *v. intr.* (-shitted, -shitting) talk nonsense; bluff. *bullshitter n.*

The bullshit that interests me falls under definition 1 of the noun, but the bullshit that interests Frankfurt is closer to what's defined by definition 2 of the noun. And that is because of the appearance of the word 'insincere' in that second definition of 'bullshit'. In definition 2 of the noun 'bullshit', bullshit is constituted as such through being the product of discourse governed by a certain state of mind. In this activity-centered definition of bullshit, the bull, conceptually speaking, wears the trousers: bullshit is bullshit because it was produced by a bullshitter, or, at any rate, by someone who was bullshitting at the time. Bullshit is, by nature, the product of bullshitting, and bullshitting, by nature, produces bullshit, and that biconditional, so understood that 'bullshitting' enjoys semantic primacy, is true of Frankfurt's view of the matter.<sup>5</sup>

Definition 1, by contrast, defines 'bullshit' without reference to the bullshit producer's state of mind. The defect of this bullshit does not derive from its provenance: almost any state of mind can emit nonsense or rubbish, with any old mix of sincerity and its lack. Here the shit wears the trousers, and *if* there are indeed "bullshitters," and "bullshittings," that correspond to the bullshit of definition 1, then they are defined by reference to bullshit: but it may be the case, as I meant to imply by that '*if*', that the words 'bullshitting' and 'bullshitter' don't have a stable place on this side of the *explicandum* divide.<sup>6</sup> However that may be, definition 1 supplies an output-centered definition of the noun: the character of the process that produces bullshit is immaterial here.

Note, moreover, how the alternatives in the brief entry on the verb 'to bullshit' match alternatives 1 and 2 in the definition of the noun (even though that entry isn't, as it perhaps should have been, sub-numbered '1' and '2'). One can "talk nonsense" with any intentions whatsoever, but one can't knowingly or inadvertently "bluff": bluffing is a way of intending to deceive. (I'm not sure, by the way, that the dictionary is right in its implication that it suffices for bullshitting, in the non-bluff sense, that

---

<sup>4</sup> 'Trivial' is very different from 'insincere', partly because it has weaker implications for the state of mind of the speaker or writer. I shall take 2 with the accent on 'insincere'.

<sup>5</sup> Frankfurt certainly believes that a person bullshits if he produces bullshit, since he thinks it a necessary condition of bullshit that it was produced with a bullshitting intention. He (in effect) raises the question whether that intention is also sufficient for bullshit at p. 9. But, although he doesn't expressly pursue that question, his definition of 'bullshit' (pp. 33-34), and its elaboration (pp. 54ff), show that he holds the sufficiency view as well. It is because Frankfurt asserts sufficiency that he can say (pp. 47-48) that a piece of bullshit can be true.

<sup>6</sup> See, further, the last two paragraphs of Section 4 below.

you produce bullshit, in sense 1: innocent producers of bullshit might be said not to be bullshitting when they produce it.<sup>7</sup>

It is a limitation of Frankfurt's article that, as we shall see, he took for granted that the bull wears the semantic trousers: he therefore focused on one kind of bullshit only, and he did not address another, equally interesting, and academically more significant, kind. Bullshit as insincere talk or writing is indeed what it is because it is the product of something like bluffing, but talking nonsense is what it is because of the character of its output, and nonsense is not nonsense because of the nonsense-talker's mental state.

### 3. Bullshit and Lying

At the beginning of his article, Frankfurt describes a complexity that afflicts the study of bullshit:

Any suggestion about what conditions are logically both necessary and sufficient for the constitution of bullshit is bound to be somewhat arbitrary. For one thing, the expression *bullshit* is often employed quite loosely—simply as a generic term of abuse, with no very specific literal meaning. For another, the phenomenon itself is so vast and amorphous that no crisp and perspicuous analysis of its concept can avoid being procrustean. Nonetheless it should be possible to say something helpful, even though it is not likely to be decisive. Even the most basic and preliminary questions about bullshit remain, after all, not only unanswered but unasked. (pp. 2-3)

I have no problem with Frankfurt's first remark, to wit, that "bullshit" has a wide use in which it covers almost any kind of intellectual fault. To circumvent this problem, to identify a worthwhile *explicandum*, we could ask what 'bullshit' denotes where the term does carry (as Frankfurt implies that it sometimes does) a (more or less) "specific literal meaning," one that differs, in particular, from the meanings carried by words that are close to 'bullshit', but instructively different in meaning from it, such as the word 'horseshit', which, at least in the United States, denotes, I believe, something characteristically produced with less deviousness than characterizes the production of (OED-2) bullshit. ...

\* \* \*

---

<sup>7</sup> See, once again, the last two paragraphs of Section 4 below.

... Frankfurt very surprisingly says, that “the essence of bullshit ... is ... lack of connection to a concern with truth— ... indifference to how things really are” (pp. 33-34), where that indifference (see the Frankfurt passage in quoted in the paragraph that follows here) is concealed by the speaker. It’s the word ‘essence’ that surprises me here: it seemed to be implied by Frankfurt’s preliminary remarks that the term ‘bullshit’, considered comprehensively, denotes no one thing whose essence one might try to specify, ... and Frankfurt had not in the interim indicated a particular *region* of bullshit, whose bullshit might, perhaps, be identified by an essence.

Frankfurt later elaborates his definition as follows:

This is the crux of the distinction between him [the bullshitter] and the liar. Both he [the bullshitter] and the liar represent themselves falsely as endeavoring to communicate the truth. The success of each depends upon deceiving us about that. But the fact about himself that the liar hides is that he is attempting to lead us away from a correct apprehension of reality; we are not to know that he wants us to believe something he supposes to be false. The fact about himself that the bullshitter hides, on the other hand, is that the truth-values of his statements are of no central interest to him; what we are not to understand is that his intention is neither to report the truth nor to conceal it. This does not mean that his speech is anarchically impulsive, but that the motive guiding and controlling it is unconcerned with how the things about which he speaks truly are. (pp. 54-55)

Notice that, when Frankfurt elaborates what is supposed to be a proposal about bullshit, he speaks not of “bullshit” but of the “bullshitter.” This confirms that it is the bull that wears Frankfurt’s trousers. But he wrongly takes for granted that that is the only important or interesting bullshit that there is.

Now, in the light of the semantic promiscuity of ‘bullshit’ that was discussed at the outset of this section, it was, so I have suggested, unwise of Frankfurt to cast his claim as one about the “essence” of bullshit, as he does in the pp. 33-34 passage. He should have submitted his indifference-to-truth thesis as an attempt to characterize (at least) one interesting kind of bullshit, whether or not there are other interesting kinds of it. Let us assess his thesis as such, that is, not with the ambitiously generalizing status that Frankfurt assigns to it, but as an attempt to characterize one kind of bullshit, and, in particular, an activity-centered kind of bullshit. I return to the distinct bullshit-*explicandum*, which corresponds to OED definition 1, in Section 4 below.

\* \* \*

#### 4. Bullshit as Unclarifiable Unclarity

Unlike Frankfurt's bullshitting, lying is identified in terms of the defect at which it aims, namely, falsehood. We clarify what a liar is by reference to falsehood, rather than the other way around; we do not, that is, when asked to characterize what falsehood is, say that falsehood is what a liar aims to say. In parallel, we might, unlike Frankfurt, seek to clarify what a bullshitter is by reference to what he aims at, to wit, bullshit. We might start with the shit, not with the bull. And that would induce us to consider OED definition 1 ("nonsense, rubbish") the one that fits the bullshit that interests me, rather than the bullshit that interests Frankfurt. My bullshit belongs to the category of *statement* or *text*. It is not primarily an activity but the result of an activity (whether or not *that* activity always qualifies as an activity of bullshitting.<sup>8</sup>)

A liar who tries to say something false may inadvertently speak the truth, whether or not he is then lying, and whether or not what he then says is a lie. And there is also the opposite case in which an honest person, by mistake, speaks falsely. The bullshit that interests me is relevantly parallel. I countenance a bullshitter who has tried, but failed, to produce bullshit—what comes out, by accident, is good sense—and I also countenance a lover of truth who utters what he does not realize is bullshit. A person may avow, in full honesty, "I'm not sure whether what I'm about to say is bullshit." These are not possibilities for the bullshit that interests Frankfurt. But they are possibilities. So the bullshit that interests Frankfurt doesn't cover the waterfront.

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.<sup>9</sup> And, oppositely, an honest person might read some bullshit that a Frankfurt-bullshitter wrote, believe it to be the truth, and affirm it. When that honest person utters bullshit, *she's* not showing a disregard for truth. So it is neither necessary nor sufficient for every kind of bullshit that it be produced by one who is informed by indifference to the truth, or, indeed, by any other distinctive intentional state.

The honest follower, or the honest confused producer of bullshit, may or may not count as a bullshitter,<sup>10</sup> but she is certainly honest, and she certainly utters (one kind of) bullshit. There exists bullshit as a feature of utterances that does not qualify as bullshit by virtue of the intentional state of

---

<sup>8</sup> See the final paragraph of this section.

<sup>9</sup> Perhaps in contrast with Frankfurt's sense, and certainly in contrast with what Frankfurt says about that sense (see pp. 47-48).

<sup>10</sup> That question is addressed in the penultimate paragraph of this section.

the utterance's producer (although that state may, of course, causally explain why the bullshit is there, and/or why what's there is bullshit).

But what *is* that feature of utterances? *One thing it can be*, at least to a first approximation, is what the OED calls it, to wit, *nonsense*. But what particularly interests me is a certain variety of nonsense, namely, that which is found in discourse that is by its nature *unclarifiable*, discourse, that is, that is not only obscure but which cannot be rendered unobscure, where any apparent success in rendering it unobscured creates something that isn't recognizable as a version of what is said. That is why it is frequently an appropriate response to a charge of bullshit to set about trying to clarify what was said. (Think of attempts to vindicate Heidegger, or Hegel. The way to show that they weren't bullshitters is not by showing that they cared about the truth, but by showing that what they said, resourcefully construed, makes sense. Those who call them bullshitters do not doubt that they cared about the truth, or, at any rate, it is not *because* of any such doubt that they think Hegel and Heidegger were bullshitters.<sup>11</sup> That Frankfurt issue isn't the issue here.)

Something is unclarifiable if and only if it cannot be made clear, but I shall not try to say what "clear" means in this essay. (I'm inclined to think it's not possible to do so, in an illuminating way.) Note, however, that there are relevantly different forms of unclarity, all of which have bearing here. There is the unclarity of a sentence itself, and then there is the unclarity as to why a certain (possibly perfectly clear) sentence is uttered in a given context: So, for example, the meaning of Wittgenstein's "If a lion could speak, we would not understand him" is in one way perfectly clear, but it might nevertheless be judged obscure, and unclarifiably obscure, by one who doubts that it carries, in context, a graspable point. There is also the unclarity of why one statement should be taken to lend credence to another statement. And there are no doubt other pertinent unclaritys too.

Note that it is not an objection to the proposed sufficient condition of bullshit that different people might, in the light of different background beliefs, impose different standards of clarity, and, therefore, identify different pieces of texts as bullshit. Some of the people might, of course, be wrong.

I emphasized "one thing it can be" three paragraphs back because defects other than unclarifiable unclarity can suffice to stigmatize a text as bullshit. I focus on this variety of the phenomenon because it commands a greater academic following than other varieties do. In the

---

<sup>11</sup> For the record, I do not believe that Hegel was a bullshitter, and I am too ignorant of the work of Heidegger to say whether or not he was a bullshitter. But I agree with my late supervisor Gilbert Ryle that Heidegger was a *shit*. I once asked Ryle whether he had continued to study Heidegger after he had written a long review of *Being and Time* that was published in *Mind*. Ryle's reply: "No, because when the Nazis came to power, Heidegger showed that he was a shit, from the heels up, and a shit from the heels up can't do good philosophy." (Experience has, alas, induced me to disagree with the stated Rylean generalization.)

various varieties of bullshit, what is wanting, speaking very generally, is an appropriate connection to truth, but not, as in Frankfurt's bullshit, as far as the state of mind of the producer is concerned, but with respect to features of the piece of text itself. Unclarifiably unclarity is one such feature. Rubbish, in the sense of arguments that are grossly deficient either in logic or in sensitivity to empirical evidence, is another. A third is irretrievably speculative comment, which is neither unclear nor wanting in logic, such as—David Miller's excellent example—"Of course, everyone spends much more time thinking about sex now than people did a hundred years ago."

I focus on unclarifiable unclarity in particular in preparation for a further inquiry into bullshit that addresses the question why so much of that particular kind of bullshit is produced in France. This kind of academic bullshit, unlike the two contrasting types of bullshit, be they academic or not, mentioned in the previous paragraph, comes close to being celebrated for its very unclarity, by some of its producers and consumers. What some of them certainly celebrate is a disconnection with truth: in what perhaps ranks as the consummation of the development of unclarity-type bullshit, a consummation that Hegel might have called "bullshit risen to consciousness if itself", truth is, in much post-modernism, *expressly* disparaged.

Although I foreswear a definition of 'clarity', I can offer a sufficient condition of unclarity. It is that adding or subtracting (if it has one) a negation sign from a text makes no difference to its level of plausibility:<sup>12</sup> no force in a statement has been grasped if its putative grasper would react no differently to its negation from how he reacts to the original statement. The deliberate bullshit published by Alan Sokal<sup>13</sup> no doubt comes out as unclarifiable unclarity, by that criterion. Note that this test does not apply to the different sorts of bullshit reviewed a couple of paragraphs back, and, being a merely sufficient condition of unclarifiable unclarity, it does characterize *all* cases of the latter either.

An objection that faces my account is that it appears to classify good poetry that isn't bullshit as bullshit, since a piece of good poetry may be unclarifiable. A tempting way of acquitting such poetry of the charge of bullshit is by reference to its designation *as* poetry, rather than *as* some sort of contribution to knowledge in a more straightforward sense. But then the same text would be bullshit or not according, Frankfurt-like, to its, as it were, intentional encasement, and I am trying to characterize an intention-independent sense of the term.

---

<sup>12</sup> This criterion of bullshit was devised by Professor Arthur J. Brown, to whom I am indebted.

<sup>13</sup> In his wonderful spoof, "Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity"—which was published as a non-spoof in the thereby self-condemning *Social Text* 46-47 (Spring-Summer, 1996), pp. 217-252.

An unclarifiable text can be valuable because of its suggestiveness: it can stimulate thought, it can be worthwhile seeking to interpret it in a spirit which tolerates multiplicity of interpretation, and which therefore denies that it means some one given thing, as a clarifiable piece of text does. So let us say, to spare good poetry, that the bullshit that concerns me is not only unclarifiable but also lacks this virtue of suggestiveness.<sup>14</sup> (I am sure that many academic bullshitters get away with a lot of bullshit because *some* of their unclarifiabilia are valuably suggestive, and therefore not bullshit. Their readers then mistakenly expect more, or most, of it to be so.)

So much by way of a preliminary attempt to identify the *bullshit* that interests me. But what reading of 'bullshitter', if any, corresponds to the bullshit that I have tried to identify? Producers of Cohen-bullshit are clearly not by nature bullshitters, in Frankfurt's sense, though Frankfurt-bullshitters often produce Cohen-bullshit, at least in the academy. Rather I would say that the word 'bullshitter' that corresponds to my bullshit has two readings. In one of its readings a bullshitter is a person who is *disposed* to bullshit: *for whatever reason*, to produce a lot of unclarifiable stuff. In a second acceptable reading of the term, a bullshitter is a person who *aims* at bullshit, however frequently or infrequently he hits his target.<sup>15</sup> (Notice that other nouns that signify that their denotations engage in a certain activity display a similar pair of readings: a killer may be a being that tends to kill, with whatever intention or lack of it (a weed-killer, for example, is a killer, and a merely careless human stomper on flowers is a (flower-) killer); or he may be a being who intends to kill, whether or not he ever does). Aim-(Cohen)-bullshitters *seek* and *rely* on unclarifiability, whereas innocent speakers of bullshit are merely victims of it. Aim-bullshitters resort to bullshit when they have reason to want what they say to be unintelligible, for example, in order to impress, or in order to give spurious support to a claim: the motives for producing bullshit vary. (And just as a person might sometimes kill, without being a killer in either of the senses I distinguished, so a person who is in neither of the senses I distinguished a bullshitter might, on occasion, produce bullshit.)

---

<sup>14</sup> I am allowing that the unclarifiable may be productively suggestive, but I would not go as far as Fung Yu-lan does: "Aphorisms, allusions, and illustrations are ... not articulate enough. Their insufficiency in articulateness is compensated for, however, by their suggestiveness. Articulateness and suggestiveness are, of course, incompatible. The more an expression is articulate, the less it is suggestive—just as the more an expression is prosaic, the less it is poetic. The sayings and writings of the Chinese philosophers are so inarticulate that their suggestiveness is almost boundless" (*A Short History of Chinese Philosophy* [New York: Macmillan, 1960], p. 12).

<sup>15</sup> Michael Otsuka comments insightfully on a familiar academic "case in which the two come apart: that is, in which someone is disposed to unclarifiable unclarity without aiming at it. Many academics (including perhaps an especially high proportion graduate students) are disposed to produce the unclarifiable unclarity that is bullshit, not because they are aiming at unclarifiable unclarity, but rather because they are aiming at profundity. Their lucid utterances are manifestly unprofound, even to them. Their clarifiable unclear utterances can be rendered manifestly not profound through clarification. But their unclarifiably unclear utterances are unmanifestly not profound. Hence it is safe for them to think that they are profound. These utterances are not profound either because they are meaningful (in some subtle way, should there be one, that is consistent with their unclarifiable unclarity) but unprofound or because they are meaningless. They are *unmanifestly* not profound because it is hard to demonstrate that they are not profound, given their unclarifiability. By aiming at profundity, these academics tend to produce obscurity. But they do not aim at obscurity, not even as a means of generating profundity" (Private communication, 2<sup>nd</sup> September, 1999).



What about the verb, 'to bullshit'? Does the producer of my bullshit, always bullshit when she produces bullshit, as Frankfurt's does? I see no reason for saying that an innocent does, especially if she's not even a disposition-bullshitter. But an aim-bullshitter who produces bullshit indeed bullshits.<sup>16</sup>

## 5. Bullshit as Product and Bullshit as Process

It matters that bullshit can come in the non-intention-freighted form by which I am exercised. For there is, today, a great deal of my kind of bullshit in certain areas of philosophical and semi-philosophical culture, and if, as we should, we are to conduct a struggle against it, the sort of struggle that, so one might say, Alan Sokal has inaugurated,<sup>17</sup> then it is important not to make false accusations, and not, therefore, for example, to charge possibly innocent traffickers in bullshit of lacking a concern for truth, or of deliberately conniving at obscurity.<sup>18</sup> Our proper polemical target is bullshit, and not bullshitters, or producers of I bullshit, as such. So while it's lots of fun, for people like me, who have a developed infantile streak, to talk about bullshit, and even just to write 'bullshit', over and over again, in an academic article, there is nevertheless, in my opinion, something important at stake here, and the character of what is at stake makes the bullshitter/bullshit distinction important.

To prevent misunderstanding, let me add that I do believe that there is quite a lot of *aiming* at obscurity in the production of philosophical bullshit, and a lot, to boot, in this region, of lack of concern with truth.<sup>19</sup> But these moral faults should not be our primary focus. For reasons of courtesy, strategy, and good evidence, we should criticize the product, which is visible, and not the process, which is not.<sup>20</sup>

---

<sup>16</sup> Let me now list some central differences between the two kinds of bullshit that I have distinguished:

	Typical Context of Utterance	Corresponding OED Definition	Primary Locus	Essence
Frankfurt's bullshit	everyday life	2	activity	indifference to truth
Cohen's bullshit	the academy	1	output	unclarifiability

<sup>17</sup> Initially in the article referenced in footnote 13, and then more comprehensively in *Intellectual Impostures*, which he wrote with Jean Bricmont (London: Profile, 1998).

<sup>18</sup> Consider this sentence from the work of Etienne Balibar: "This is precisely the first meaning we can give to the idea of dialectic: a logic or form of explanation specifically adapted to the determinant intervention of class struggle in the very fabric of history" (*The Philosophy of Marx* [New York: Verso, 1995]). If you read that sentence quickly, it can sound pretty good. The remedy is to read it more slowly, and you will then recognize it to be a wonderful paradigm of bullshit: yet I know Balibar to be an honest thinker.

<sup>19</sup> The evidence assembled in Sokal and Bricmont's *Intellectual Impostures* proves, so I think, the truth of those beliefs.

<sup>20</sup> We may hope that success in discrediting the product will contribute to extinguishing the process. I try to contribute to the project of discrediting the product in an unpublished and unpublishable discussion of "Why One Kind of Bullshit Flourishes in France," a draft of which will be supplied upon application to me.