Bullshit as counterfeit evidence

Abstract

Frankfurt's idea that bullshit is 'fake', 'counterfeit', and 'phony' has been overshadowed in the subsequent debate by his claim that the bullshitter is indifferent to truth. The former insight deserves to be taken more seriously. I begin by asking whether bullshit is a fake assertion and argue that it isn't. What makes it bullshit, rather, is that the speaker's performance is meant to raise the audience's credence in some proposition *h* by resembling genuine evidence for *h*, while simultaneously concealing that it lacks the features of genuine evidence. Bullshit, on this analysis, is best understood as a communicative act that the speaker intends to serve as 'counterfeit evidence'. I then use the ideas of relevance theory to distinguish bullshit from plain lying. Since deception is a corollary of counterfeiture, it is also an intrinsic element of bullshit. This helps to contrast the present view with other accounts in the literature.

Keywords: bullshit, assertion, fakery, counterfeiture.

1. Already at the outset of the modern debate Max Black aligned bullshit with 'perversions of verbal interaction, animated by deliberate deceit' and with 'deliberate falsifications of the constitutive framework' of the verbal interaction (1983:131).[1] In his eventual definition of bullshit, however, Black made little use of these observations and said simply that it is a 'deceptive misrepresentation [...] of somebody's own thoughts, feelings, or attitudes' (1983:143).

Deception remained central to Harry Frankfurt's more influential treatment, where the bullshitter's distinctive feature was said to be the concealment of his indifference to truth (1988:130). Most subsequent critics have pressed Frankfurt on the truth-indifference condition itself. Should bullshitters be necessarily indifferent to truth? Should such indifference, when present, be concealed?[2]

I put this line aside here in order to focus on a different strand in Frankfurt's account, the intuition that bullshit is 'fake', 'counterfeit', and 'phony' (1988:128). I argue that bullshit is best understood as the production of *counterfeit evidence*. Deception, on this account, is a corollary of counterfeiture, not an independent condition.

I begin in §2 by drawing on 'dual content semantics' to characterise fakery and counterfeiture for artefacts. In §3 I consider whether bullshit is a fake assertion. Having argued that it's not, in §4 I suggest that bullshit is instead the production of counterfeit evidence. In §5 I argue that this initial formulation is too permissive, since it can't distinguish between bullshit and plain lies. I revise it using the relevance-theoretic notion of an informative intention and a new notion of an *indicator*. In §6 I extract the explications that follow, noting where they depart from Black and Frankfurt on the object of the bullshitter's deception. In §7 I test the account against further putative cases from the literature: some are vindicated as bullshit on sharper grounds than their original authors offered, and some, I argue, should not have been classified as bullshit in the first place.

2. To get a grip on the ideas of fakery and counterfeiture, here's an informal paraphrase of some claims of the dual content semantics, or DCS (Del Pinal, 2015, 2018). Thinking of artefacts involves associating with them characteristics like prototypical appearance, materials, and function: a gun is something with a gun-like appearance, made of gun-like parts, meant to shoot. These 'compositional' properties (C-properties) are part of the meaning of 'gun' and instrumental for competence in using it, but they do not fix its extension. The extensional property (E-property) of the common noun 'gun' is just the atomic property of being a gun.

---

1 Black spoke of 'humbug'. I treat it as synonymous with 'bullshit'.

2 Truth indifference: e.g., Stokke (2018), Carson (2016). I address this debate elsewhere [further details omitted for anonymity].

A ‘ fake gun ’ is not a gun with certain characteristics. It isn’t a gun at all. But not every non-gun is a fake gun. Fake guns are made to look like guns but not to shoot. In the idiom of DCS, some C-properties of ‘ gun ’, like its function, are ‘ uploaded ’ to the E-property of ‘ fake gun ’. The essential properties of a fake artefact thus include intended inability to perform the genuine function and intended perceptual similarity to the genuine item.

Consider next ‘ counterfeit *N* ’. Like a fake *N*, a counterfeit *N* is a non-*N* designed to look like *N*. Unlike a fake *N*, a counterfeit *N* *is* designed to be used or to function as an *N*. ‘ Forged *N* ’ is a closely related idiom whose E-property likewise refers to a corresponding act of deception.

English usage of ‘ fake ’ is indiscriminate enough that the fake/counterfeit distinction is often blurred, especially for entities with a social role.[3] There are ‘ fake smiles ’ designed to fulfil the social function of the real ones, but no ‘ counterfeit smiles ’. There are ‘ fake lawyers ’, but no ‘ counterfeit lawyers ’.[4] So as not to butcher the familiar English usage, I will here stipulate:

‘ fake *N* ’ := ‘ counterfeit *N* ’.

The underlying contrast is intact. I simply reserve both ‘ fake ’ and ‘ counterfeit ’ for the deception-involving case: an artefact or act intended to be a non-*N*, to resemble *N*, and to function as *N* does.

3. Let’s now look at Frankfurt’s central case of bullshit (call the protagonist ‘ Bullmann ’):

(1) Bullmann’s oration: A Fourth of July orator goes on bombastically about ‘ our great and blessed country, whose Founding Fathers under divine guidance created a new beginning for mankind. ’ But the audience’s opinion about American history does not motivate his speech. Instead, the orator wants his audience to walk away with a certain impression of his character: that he is a patriot, a man of depth, someone who feels the weight of history and bows humbly before God. (Frankfurt, 1988:121)

So Bullmann *is* faking something, we agree with Frankfurt, but what? In the first place, he is not faking an English *sentence*, a locution: he is a competent English speaker uttering English sentences in the ordinary way. As already noted (§1), Frankfurt’s own suggestion, that he is faking his concern with truth, doesn’t work: there are just too many problems with this notion. There is another idea in the close vicinity, though. The object of Bullmann’s fakery, if there is one, must be a pragmatic category. Assertion is the natural candidate: bullshit is a fake assertion, or equivalently, when a bullshitter utters a declarative sentence and appears to assert *p*, he in fact fakes the assertion of *p*.[5] Thus let’s unpack (1):

(2) *P* ≈ ‘ The Founding Fathers defended freedom. ’
Q ≈ ‘ I am a thinker and a patriot. ’

When Bullmann utters the English sentence (2), he non-asserts the proposition that the Founding Fathers defended freedom. Well, does he? It’s not possible to answer in full generality, as there are quite a few competing accounts of assertion. Let me explore several possibilities.

---

[3] The contrast stays explicit in German:

‘ fake gun ’ = ‘ unechte Waffe ’
‘ fake smile ’ = ‘ falsches Lächeln ’
‘ fake lawyer ’ = ‘ falscher Anwalt ’
‘ counterfeit coin ’ = ‘ gefälschte Münze ’
‘ forged Titian ’ = ‘ gefälschter Tizian ’.

[4] Guerrini (2024:179, 189) has recently taken this as grounds for collapsing ‘ fake ’, ‘ counterfeit ’, and ‘ forged ’. I read it instead as evidence that ‘ fake ’ in English is polysemous.

[5] Bullshit as ‘ fake ’ (=‘ counterfeit ’ in our terms) assertion: Kotzee (2020).

Some theorists take assertion to be a vehicle for expressing the belief that p.[6] Thus in uttering (2) Bullmann should fail to express his belief that the Founding Fathers defended freedom. This seems false. For consider the following exchange:

(3) {*Context*: Bullmann the orator delivers a speech containing (2). The auditors quietly suspect that he is a bullshitter.}

a. (Auditors:) — Well, so you didn't express your belief that the Founding Fathers defended freedom. That's clearly so, since you don't believe this in the first place.

b. (Bullmann indignantly:) — But of course I did. How *else* could I have expressed it? And for the record, I do believe that the Founding Fathers defended freedom.

As soon as we grant that Bullmann may well believe P, and that Bullmann's defence is not vacuously true (which it would have been if he didn't believe P), the auditors' complaint is a *non sequitur*. And I think we must. There is no reason why Bullmann can't believe that the Founding Fathers defended freedom, and why he shouldn't consider himself an American patriot. Might we require further that he be able to justify his belief? or that he know the proposition the Founding Fathers defended freedom? Very well, he can do that too. There is no reason why Bullmann can't be a tenured professor of American history. Generally, Bullmann need not have any special epistemic disability vis-à-vis P.

An even more straightforward reasoning holds for the view that assertion is a vehicle for presenting propositions as true.[7] However you think of 'presenting-as-true', Bullmann is doing just that. He is clearly not presenting it as false, or as open to discussion, or as asserted by some other speaker. On both the belief view and the truth view, utterances of declarative sentences in a play should be the natural candidates for merely fake assertions. Lies, on the other hand, will be counterfeit assertions. When you lie to me by asserting p, you are aware that your utterance is not an expression of your belief (since you believe $\sim p$), but you want me to treat it as a genuine expression of your belief that p. Analogously for truth presentation. Bullshit is neither this, nor that. You can bullshit by lying, but also by telling what you genuinely believe. An actor on stage can bullshit too, but whether he does so is not entailed by the fact that he utters sentences from Molière.

I've just complained that generalisations are dubious, but here is a speculation: If assertion is just a relation r between the speaker and the proposition, then for every account of r, the bullshitter s, on some occasions, is r-related to the proposition p his utterance is taken to express, just like a manifest non-bullshitter s' is r-related to *his* proposition p'. Hence the bullshitter's utterance will count as a genuine assertion.

What of the overtly pragmatic views of assertion? To assert p, on this view, is to interact with the audience and, generally, with the context in a certain way. The most familiar version was defended by Stalnaker.[8] To assert p is to 'propose' to update the common ground by shrinking the context set and throwing out the worlds where p is false. It is an attempt to evolve the conversation so as to get the asserted content p to be presupposed by the participants. Well, if assertion is so defined in terms of its conversational effects, then there is something defective about Bullmann's utterance. Arguably he is not trying to adjust the common ground. As Frankfurt imagines the situation, the auditors already accept P and the rest of the propositions

6 The belief view of assertion: e.g., Bach and Harnish (1979), Green (2009), Siebel (2020).

7 The truth view of assertion: e.g., Wright (2003:29). This, I think, is also the more recent view in Pagin (2020, 2011). Attempts to explicate 'presenting-as-true': Searle (1986:217), Pagin (2011:123ff), Marsili (2024:549), Marsili and Green (2021:23). Some of these attempts are relevant for the theory of bullshit, but there is no space to address them here.

8 Pragmatic views: e.g., Stalnaker (2014:51ff) and Clapp (2020).

in his speech, perhaps with different degrees of confidence or commitment (and that's all that counts for context changes elicited by assertions). Even if some of them don't, Bullmann is not in the business of convincing them otherwise. If he were, if his primary goal were to persuade the possibly sceptical audience of American greatness, we would not likely have judged him a bullshitter (neither would Frankfurt).

In other words, Bullmann is guilty of violating the maxim of quantity, or equivalently Stalnaker's first principle of assertion (1999:88). And this is so not by mistake, but by design. The speaker doesn't even try to evolve the common ground. He is aware that his utterance should fail to do so, but he makes it all the same. Thus, adopting Stalnaker's view, we have:

(4) An utterance u made by s and expressing p is a fake assertion.
= u is a non-assertion, s intends u to look like a genuine assertion (by using a declarative sentence, appropriate intonation, gestures, facial expressions), s does not intend u to be used in updating the context by removing $\sim p$-worlds from the context set.

To assert P in (2) is to assert a platitude. Not every expression of platitudes counts as a fake assertion, however. Contexts of politeness and small talk are a primary example here. If I say, 'Nice weather, innit?', it may be clear from my intonation and my mien that I don't really intend to evolve the common ground and draw your attention to the weather outside. I don't intend my utterance to look like a genuine (indirect) assertion. Moreover, my interlocutors should misunderstand the context to believe that I do. Hence small talk consists of mere non-assertions, rather than fake assertions. This is in contrast to assertoric utterances in a Molière performance that are supposed to resemble as much as possible real assertions. Such utterances are the theatrical analogue of a fake gun in the original DCS sense, i.e. non-assertions designed to resemble assertions, but without the deceptive intent that would make them counterfeit.

On the present view, Bullmann is in a similar position. Festivities are not occasions for revelations and debates: they are designed for platitudes. Thus, had we been told that Bullmann made a 'perfunctory' speech in a dull, barely audible voice, his utterance (2) would be just a non-assertion. Instead, Frankfurt insists, his speech is 'bombastic'. Bullmann plainly tries to make his utterance look like a genuine assertion. For this reason, and because its content remains a platitude, it is a fake assertion. In saying that we assume that the assertoric act has certain perceptual marks like the declarative sentence used to make the putative assertion, along with typical paralinguistic elements. The assertion itself is defined by its effect on the common ground: namely, its disposition to update the common ground.

As a general account of bullshit, this can't be right, however: the failure to update the common ground is not its essential feature, and Bullmann's speech is only incidentally implicated in such a failure. The reason is plain: bullshitters often do intend their interlocutors to believe the proposition their utterance expresses and do not think of that proposition as a platitude already in the common ground.

Sometimes they lie.[9] When the confidence trickster Frank Abagnale in *Catch Me if You Can* is caught red-handed by the FBI agent Carl Hanratty, he coolly decides to bullshit his way through and pretends to be a USSS agent.[10] As part of that bullshit routine, he says, 'My name's Barry Allen, United States Secret Service.' Abagnale is lying *and* bullshitting: his lie is also bullshit. Plainly his utterance is also an assertion intended to update the common ground of the conversation with Hanratty.

Other times, they tell the truth. It's easy to imagine that Abagnale was in fact a receptionist on the USSS payroll, and that his name was unknown to Hanratty. He could then truthfully say,

[9] Bullshit is compatible with lying: e.g., Carson (2016).

[10] The Abagnale scene: https://www.youtube.com/watch?v=gFr0_ywVdhY

‘My name’s Abagnale, United States Secret Service.’ In that context, his utterance could be regarded as bullshit, too. Of course, it would again be a genuine assertion.

I conclude that the failure to update the common ground is not a distinctive feature of bullshit, so that also according to Stalnaker’s common-ground view, bullshit utterances are not necessarily, not even typically, non-assertions. Hence they should count neither as fake, nor as counterfeit assertions.

4. If the reasoning of the previous section is on the right track, bullshit cannot be characterised as a dubious speech act. To do justice to Frankfurt’s descriptions of it as ‘deception’ and ‘phoniness’, we must try again.

Let’s return to gunmaking one more time. Gunnar wants you to believe that he is a top gunsmith, so he invites you to his workshop. It is common knowledge between the two of you that top gunsmiths work in a particularly elegant, fluent way. That is indeed how Gunnar normally works. Today, however, he is busy, so he arranges for his twin Ragnar, another top gunsmith, to be in the workshop when you arrive. Ragnar is not privy to the ploy and unaware of your presence. He is really producing a real gun. You observe Ragnar, come to believe that he is Gunnar, and on the basis of this false belief conclude that Gunnar is a top gunsmith.

In this scenario there is no counterfeit object, no counterfeit action. There is, however, counterfeit *evidence* created by Gunnar:

(5) E: Gunnar is working fluently in the workshop.

E': Ragnar is working fluently in the workshop.

H: Gunnar is a top gunsmith.

E is evidence for H whereas E' is not. On its own E' is no more counterfeit evidence for H than a certain bag, on its own, is a counterfeit Chanel bag. Counterfeiture requires intentional design, which is what we have here. Gunnar *arranges* for E' to be observed by you, and he does so because E' perceptually resembles E in your circumstances. In general:

(6) An event e' is counterfeit evidence for h.

= there is an event $e \neq e'$, an agent a, an observer o, a circumstance c such that: (i) e increases o’s credence in h, (ii) e' does not increase o’s credence in h, (iii) the perception of e' in c increases o’s credence in the occurrence of e, and (iv) a provides e' to o in c with the purpose of increasing o’s credence in h.[11]

This formula entails that deception is part of counterfeiture. Once o is aware of the ploy to provide him with e', he must become aware that he perceives e' rather than e and that he has no additional ground to believe h. To prevent that, a must keep his ploy hidden from o.

Now recast Bullmann in terms of (6). Just like Ragnar, Bullmann may really be a patriot. His general behaviour and biography may be sufficient for judging him one. On this particular occasion, nevertheless, he is manufacturing counterfeit evidence. As Frankfurt puts it, ‘what he cares about is what people think of him. He wants them to think of him as a patriot’ (1988:121). Bullmann asserts P only to *appear* a patriot, that is, he presents evidence to his audience with an ulterior purpose.

It is essential to that purpose that it remain hidden. Suppose Bullmann prefaced his speech by saying, ‘Though my mind is on other things, I must now say a few words so as to appear a patriot.’ The audience would be wrong to infer that Bullmann is *not* a patriot, since for all they know he may be. But his speech could no longer be good evidence of his patriotism (though

[11] We can assume the transitivity of evidence, taking e to screen off the perception of e' from h.

it is evidence of his willingness to conform to convention). The reasoning is straightforward. Patriotism does not consist in delivering speeches. The speech counts as evidence of patriotism only on the assumption that the speaker is passionate about what he says and that this passion explains why he says it. Once Bullmann's ulterior purpose is recognised, that assumption drops, and with it the speech's status as evidence of patriotism. We thus have:

(7) E: Bullmann uttering P out of conviction.

E': Bullmann uttering P only to impress the audience.

H: Bullmann is an American patriot.

On a simple Bayesian interpretation, let the audience's prior be $\mathrm{P}(H) \approx 1/2$. Then $\mathrm{P}(H \mid E) \gg \mathrm{P}(H)$, since $\mathrm{P}(E \mid H) \gg 0$ and $\mathrm{P}(E \mid {\sim}H) \approx 0$. On the other hand $\mathrm{P}(H \mid E') \approx \mathrm{P}(H)$, since $\mathrm{P}(E' \mid H) \approx \mathrm{P}(E' \mid {\sim}H)$. That is, an utterance made purely to impress is roughly equally probable whether or not the speaker is a patriot. Hence, on its own, E' is not evidence for H. Letting r be the audience's sensory representation of E' in c, we can say: Bullmann's ploy is to ensure that $\mathrm{P}(H \mid r) \gg \mathrm{P}(H)$. As the audience cannot distinguish r from a sensory representation of E, it updates as if E had occurred.

The reasoning generalises, subject to one proviso. What counts as counterfeit evidence depends on what we take genuine evidence to be. Bullmann's speech counts as evidence of patriotism only on the assumption that his passion is a genuine motive explaining why he speaks as he does. Not every speech made with an ulterior purpose involves such an assumption. Suppose you lecture on *Philebus* with the purpose of impressing the audience with your scholarship. Suppose this purpose is somehow evident to the audience. You are not thereby a bullshitter. A lecture counts as evidence of scholarship on the assumption of competence, which is consistent with aiming to show off. If you *did* fail the conditions that make the lecture evidence of scholarship, by mechanically memorising its content or having it generated by AI, *then* you would be bullshitting your audience in the way Bullmann does. Then you would also have to hide that failure for the lecture to preserve its evidential role.

5. The explication (6) casts the net too wide, however. We want the production of counterfeit evidence to be characteristic of bullshit, to the exclusion of other kinds of deception. And (6) does exclude some. Deception by direct neural intervention is ruled out trivially.[12] If Gunnar inserts a chip in your brain so that you come to believe he is a master gunsmith, he has deceived you without producing counterfeit evidence, since no evidence has been arranged at all.

Lies are a different matter. Consider a 'plain lie' to be contrasted with an Abagnale lie soaked in bullshit. A plain lie is intended to be accepted only in virtue of its propositional content. It is Monday, and I ask you what day it is. You know it is Monday, but you lie by saying, 'Today is Sunday.' We have:

(8) E: (You truthfully utter:) 'Today is Sunday.'

E': (You utter a lie:) 'Today is Sunday.'

H: It is Sunday.

While E and E' are distinct events (one an instance of truth-telling, the other of lying), they are perceptually indistinguishable. E raises my credence in H. E' does not: on its own, a lie is no evidence for its content, and were I aware of the lie my credence would not rise. My perceiving E' nevertheless raises my credence in E's occurrence, since from the utterance alone I can't tell

[12] I follow Carson (2010:48) in counting such cases as deception. For an opposite view see Fallis (2024:226ff).

the two apart. You provide E' with the purpose of raising my credence in H. Every condition of (6) is satisfied, and the plain lie counts as counterfeit evidence. That's an unwelcome result.

For a revision, let's go back to the basics. Consider a counterfeit Chanel bag. Its maker does not merely slap on it the label 'Chanel', still less the sentence 'This is a Chanel bag.' He labours to replicate the visual appearance, the materials, the stitching, the zipper. What will dupe gullible observers is the bag's perceptual qualities, not a mere testimony. A plain liar operates differently. His utterance is, of course, also a form of sensory evidence. But its perceptual features are only the means by which you recover what he said. He doesn't intend them to be reasons for accepting what he says. His lie, *qua* lie, is supposed to have the same effect whether delivered in person, in a paper note, or in a WhatsApp message.

The contrast is between two kinds of evidence. The liar intentionally conveys p (the proposition his utterance truth-conditionally expresses or conversationally implicates), and at least one of the reasons why the observer is supposed to accept p is her grasp of that intention. The counterfeiter intentionally conveys q (distinct from the proposition his utterance expresses), but the reason the observer is supposed to accept q is *not* her grasp of that intention.

It's useful to unpack the last claim in the terms of relevance theory.[13] The Chanel counterfeiter has an informative intention of the form, 'The customer should note the similarity between my product and a genuine Chanel bag.' Yet he does not intend the customer to recognise *that* intention, let alone to believe the bag genuine *for that reason*. Same for Bullmann, who does not intend his audience to believe him a patriot by recognising his informative intention, 'The audience should note my passionate delivery, my choice of words etc.'

The bullshitter may certainly have *other* informative intentions that he does intend the audience to recognise, and various communicative intentions besides. But the propositions carried by these other intentions are epistemically independent of the hypothesis h ('This bag is genuine', 'I am a patriot'). The reasons for the growing credence in h are not supposed to include metarepresentations of those other intentions. We should therefore impose a constraint on the kind of event meant to serve as counterfeit evidence:

(9) An event e is an `indicator` of h for an observer o.[14]
= e increases o's credence in h, but not in virtue of o's metarepresenting any informative intention of a such that o's increased credence in h depends on that metarepresentation.

The revision of (6) takes the following shape:

(10) An event e' is `counterfeit evidence` for h.
= There is an event $e \neq e'$, an agent a, an observer o, a circumstance c such that: (i) e is an indicator of h for o, (ii) e' is not an indicator of h for o, (iii) the perception of e' in c increases o's credence in the occurrence of e, and (iv) a provides e' to o in c with the purpose of increasing o's credence in h.

Thanks to the revision, we may distinguish lies from bullshit. Yet the same utterance may be both a lie and a bullshit. A speaker may wish to instil in you a false belief by two routes, by asserting p *and* by fabricating perceptual indicators that raise the likelihood of your accepting p. Then his utterance won't be a *plain* lie, as it is no longer intended to be accepted only in virtue of its propositional content.

---

[13] Relevance theory, informative and communicative intentions, metarepresentation: Sperber and Wilson (2026:317ff), Wilson and Sperber (2012:150ff).

[14] I borrow the 'indicator' terminology from Dretske (1988:54ff). I can't detail here further connections to Dretske's discussion.

This is Abagnale's case (§3). Had he confined himself to saying 'I am a USSS agent', he would have told a single, and probably unsuccessful, lie. In the event, he created high-quality counterfeit evidence of being an agent:

(11) E: A real USSS agent with a characteristic appearance, producing legitimate credentials, describing a collection of confiscated cheques, the suspect's typewriter and so on.

E': Frank Abagnale speaking confidently, producing an empty wallet, describing his own tools of forgery.

H: This individual is a USSS agent (and not a forger).

A Bayesian interpretation parallels that of (7). Hanratty's prior is $\mathrm{P}(H) \approx 0$. E would increase it sharply, E' on its own wouldn't. However, the sensory representation r of E' in c is indistinguishable from a representation of E in c. Hence, $\mathrm{P}(H \mid r) \gg \mathrm{P}(H)$.

6. At least the way Frankfurt intended it to be, bullshit turns out to be an act of creating counterfeit evidence. The bullshitter makes an utterance u that expresses a proposition p. What is at stake for him, what he wants the audience to believe, is a proposition q that in general is logically independent of p. His motivation in uttering u must be hidden from the audience. By hearing u the audience will likely believe q. But were it aware that u was made only or chiefly for that purpose, it would be unlikely to come to believe q. The bullshitter's utterance, so described, is counterfeit evidence for q. Hence the following explications:

(12) 'x bullshits y.'
= 'x provides counterfeit evidence to y.'
'x bullshits y by asserting p.'
= 'x intends the assertion of p to serve as counterfeit evidence presented to y.'
'x bullshits y by ϕ-ing.'
= 'x intends ϕ-ing to serve as counterfeit evidence presented to y.'

In contrast to fakery, there is no counterfeiture without deception. Hence we can accommodate Frankfurt's claim, inherited from Black, that deception is intrinsic to bullshit.

We dispute, however, Black's idea that the object of the bullshitter's deception is necessarily the speaker's internal state, his 'thoughts, feelings, or attitudes' (1983:143). The proper object of deception is just the fact of counterfeiture. The bullshitter's ϕ-ing, such as asserting p, appears to be evidence for some h when in reality it is not, and that is what the bullshitter wishes to hide. The identity of h is undetermined. It can be anything at all, though it may often be concerned specifically with the speaker's mental states.

We similarly dispute Frankfurt's own understanding of the object of that deception, the claim that the bullshitter hides his indifference to the truth of p. On many occasions the speaker is not hiding such indifference, since, as Frankfurt's critics have pointed out, he might not be indifferent to p's truth to begin with. Still there is a grain of truth in Frankfurt's intuition. Indifference to the semantic or epistemic status of p is often a collateral condition of bullshit, since the bullshitter does not assert p to insist on its truth or to inform us that p. His assertion serves some other goal, one in principle having nothing to do with p. Hence in many instances he would be just as happy to assert some p' logically and epistemically independent of p.

7. I'll close by looking at some further putative examples of bullshit suggested in the literature. I'll argue that they fit the account above rather well. Some it classifies as bullshit on sharper grounds than the original author offered. Some it plausibly declines to classify as bullshit. Some cases admit multiple interpretations, and our account can discriminate between them.

Student's exam (Carson, 2016). A student X must answer an exam question. She knows nothing about the question and can't construct any meaningful argument. Rather than leaving the answer blank, she writes several paragraphs of plausible-sounding but wholly irrelevant material that never addresses the actual question. She does this because, according to the course requirements, even a completely off-topic answer earns 60%, while blank space earns zero. Since this provision is common knowledge between her and the instructor, X doesn't intend to deceive the instructor about what she is doing. This, Carson argues, is a case of 'open transparent bullshiting in which the bullshitter has no hope or intention of deceiving anyone' (2016:58). As such, it undermines our results in (12). Since counterfeit evidence is meant to deceive, bullshit can't be explicated as a form of counterfeit evidence.

But I am not convinced. Supposing that the 60% rule is common knowledge between X and her instructor, X has no need to deceive about *whether* her answer addresses the question. The rule itself, however, is unrealistic. Real syllabi do not reward the submission of *any* non-blank prose: soft grading rules typically reward something like demonstration of engagement with, or knowledge of, the subject matter. On the realistic reading, X's task is to demonstrate knowledge she doesn't have. This is confirmed by the particulars of Carson's scenario when the student is writing about Henry Ford on a question about Dodge vs. Ford. Then it is a straightforward case of counterfeit evidence: we have an event fine-tuned to increase the instructor's credence in X's knowing the subject. This is done by mimicking, however ineptly, the genuine evidence of such knowledge.

But let us now take Carson's stipulation at face value. Even then we can explain why X still tries to bullshit the instructor. Consider first what bare compliance with the 60% rule should look like. Make the rule explicit:

(13) The student will receive 60% provided she submits *any* fragment of grammatical English prose.

There might be a student who complies with (13) by merely writing half a page on LA weather, and who would not intend to be seen as doing anything more than that. She would just present genuine evidence of grammatical English prose. No counterfeiture would be involved in this act of bare compliance.

The student X acts differently, however, when she writes on economics and Henry Ford. Her output is likely meant to be taken as a partial demonstration of knowledge of the subject, conceivably in the hope of recovering something beyond the 60%. The essay is tailored to be perceived as a cue of her knowing the subject: it is topic-adjacent, plausible-sounding, structured as an answer rather than as filler. Had the topic-adjacency been a natural byproduct of her attempt at compliance, it could have functioned as just such a cue. Since it is instead 'engineered' to be so understood, it is counterfeit in the sense of (10). What X must conceal is just the engineering: if the instructor came to see the topic-adjacency as strategic rather than natural, the essay wouldn't raise his credence in the student's knowing the subject. It would instead be understood as evidence of the student's resourcefulness in exploiting (13). The transparency of the rule, in other words, is compatible with the opacity of X's fine-tuning. Our account locates counterfeiture and bullshit in the latter.

Incidentally, in this regard X's case runs parallel to Bullmann's. The conventions of Fourth of July oratory are common knowledge to his audience as much as the rule (13) is common knowledge to her instructor. Hidden in each case is whether *this* performance (speech, essay) has the motivation that would make it genuine evidence (patriotism, knowledge of the subject) or merely the desire to be taken to have it.

Cato the Younger (Sorensen, 2011). Cato the Younger filibustered the Senate with rambling and irrelevant speeches to run out the clock until dusk, when the Senate was required to adjourn. He was happy to violate the maxims of quantity, relation, and manner, just not the quality: as a Stoic committed to honesty, Cato did not lie. Instead, Sorensen concludes, he bullshitted (2011:406).

Well, did he? Here, the disagreement is whether deception is necessary for bullshit. Now Cato's filibuster admits two readings. On the more likely one, it is transparent. The audience knew he was only talking to consume time. No-one was misled into thinking that this was a substantive contribution. Conceivably Cato was aware of all this and made no attempt to mislead (indeed, such an attempt wouldn't be Stoic!). His speech was meant to be recognised as mere verbiage. In this respect it is similar to a student 'barely complying' with (13): both only aim to comply with a formal regulation, and their audience knows that, and they know their audience knows that. There is no sense here in saying, 'Cato bullshitted his audience (say, into believing p).' Our account predicts exactly that.

A non-transparent variation is also possible, however. Suppose Cato delivers the speech in a manner that leaves the audience uncertain whether his contribution is substantive or merely procedural (i.e. whether he is filibustering). Then let:

(14) H: 'Cato is making a substantive contribution to the debate.'

E: 'A senator delivering a substantive (possibly verbose) speech on the matter.'

E': 'Cato delivering a long, content-thin speech aimed at consuming time.'

r: The audience's sensory representation of E' on the occasion c.

In the transparent variation, $\mathrm{P}(E \mid r) \approx 0$, hence $\mathrm{P}(H \mid r) \approx \mathrm{P}(H)$, and the condition (iii) of (10) fails. In the non-transparent variation, $\mathrm{P}(E \mid r) > 0$, hence $\mathrm{P}(H \mid r) > \mathrm{P}(H)$, and (iii) is satisfied. The condition (iv), that Cato provides E' with the purpose of raising the audience's credence in H, is satisfied just in case the chosen non-transparency reflects an intention to be so received. On the natural reading of chosen non-transparency, it does. The non-transparent variation is, therefore, bullshit on exactly the grounds that make Bullmann's speech bullshit.

We can now pinpoint the disagreement with Sorensen's view a bit more clearly. Sorensen's view locates bullshit in both variations, as Cato violated the same maxims in both. The account above locates it only in the non-transparent variation. This is not just a verbal issue. For if the transparent Cato bullshits in just the sense that Bullmann does, then the deceptive mimicry of evidence and the open flouting of conversational norms are collapsed into a single category. Then we need a further term to mark the difference between them. Our account requires no such detour. Bullshit is the deceptive performance that presents itself as evidence of something it is not. Hence the transparent Cato, whatever else he is guilty of, is not guilty of that.

Student's excuse (Wreen, 2013). A student Y approaches his lecturer at the end of term seeking special dispensation. He claims that mid-semester his father suffered a stroke, and his mother broke her hip. His brother, who might otherwise have helped, had been called up for military service. The student requests a delayed exam and alternative coursework, promising to complete everything within six months.

This is meant to be another case of a lie that is also bullshit. Each individual proposition expressed in the student's narration may be a lie, and the lecturer is meant to accept the propositions on the basis of Y's testimony. But Y does not rely on testimony alone. The student likely counterfeits the 'texture' of the narration, rich with paralinguistic elements like intonation and emotional colouring. Unlike a plain liar, Y doesn't intend the lecturer to comply with his request only on the basis of his testimony, i.e. only because of grasping any of his informative or communicative intentions. He certainly wants the lecturer to swallow the lie, but he first

‘ presents ’ himself and ‘ appears as ’ a victim of circumstances. Thus let H = ‘ Y is the victim of these circumstances ’, E = ‘ Someone genuinely in such circumstances recounting them ’, E' = ‘ Y recounting them as he does. ’ Y’s narration is an indicator of H in the sense of (9): it raises credence in H not in virtue of the lecturer’s metarepresenting Y’s informative intention but in virtue of the perceptual features of the utterance token. And since Y is not in such circumstances and knows he is not, the indicator is counterfeit: his narration is meant to be the counterfeit evidence of victimhood.

As with the Abagnale case, here too we see the utility of the condition (9). Wreen’s verdict that Y is bullshitting is vindicated, but on grounds that discriminate Y’s narration from a plain lie where the lecturer would accept the propositions only on the basis of testimony.

Wishful thinker (Stokke, 2018). Jack and Julia are heading to Chicago for a Cubs game. Julia, a Cubs fan, hopes for dry weather. When Jack mentions hoping it won’t rain, Julia replies with a confident air, ‘ This time of year, it’s always dry in Chicago. ’ But she has no basis for her claim.

Andreas Stokke argues that Julia is bullshitting because, having no good reasons for her assertion, she is indifferent toward contributing reasonable, justified answers to the QUD (2018:155). But is such indifference sufficient for bullshit? On the account above, it isn’t: bullshit requires the production of counterfeit evidence, and counterfeit evidence requires deception. Whether Julia bullshits depends on whether she is being deceptive. We can thus imagine two versions of the event.

One version is parallel to Wreen’s student Y. Julia wanted Jack to believe P = ‘ Chicago is dry etc. ’ But first, she wanted Jack to believe Q = ‘ Julia is confident about P’s truth. ’ Her confident air was intended as an evidential indicator for Q. But since she had no good reasons for being confident (and knew that too), her air was counterfeit. She bullshitted Jack into believing Q, hence indirectly also into believing P. Here our verdict converges with Stokke’s, but on different grounds: what makes the case bullshit is not Julia’s indifference to the QUD, but her engineering of a perceptual indicator to raise Jack’s credence without any good warrant.

In the other version, Julia is doing nothing of the sort. She just chimes in to manage the conversation: to stop Jack worrying, to close the topic. Or perhaps she speaks without thinking. Her confident air is not meant to be an indicator of anything in particular. It’s just a form of casual speech, not a contrived signal. Since none of the conditions of (10) are satisfied, there is no bullshit.

In the latter scenario, our account yields a result very different from Stokke’s. Julia may be epistemically irresponsible, careless about the QUD, and unjustified in her assertion. Yet, because she doesn’t deceive Jack, she doesn’t bullshit him, either. This keeps the concept of bullshit fine-grained. On Stokke’s view, if you chime in politely, or fill awkward silence with unwarranted remarks, or repeat things half-remembered—any of that will make you a bullshitter like Bullmann. This groups together communicative behaviours that are fundamentally distinct: Bullmann the bullshitter *par excellence* fabricates evidence of a motivation he lacks, the others merely assert without reasons (which is possible, e.g., on the truth-presentation view—see §3). Both may be criticised, but not in the same way and not for the same reason. ‘ Bullshit, ’ in Frankfurt’s and Black’s hands, was a term for the first kind of wrongdoing, a form of communicative dishonesty. Failures of epistemic diligence, when not dishonest, should better be regarded as a different category.

References

Bach, K. and Harnish, R. M. (1979). *Linguistic Communication and Speech Acts*. MIT Press, Cambridge, MA.

Black, M. (1983). The prevalence of humbug. In *The Prevalence of Humbug and Other Essays*. Cornell University Press, Ithaca, NY.

Carson, T. L. (2010). *Lying and Deception*. Oxford University Press, Oxford.

Carson, T. L. (2016). Frankfurt and Cohen on bullshit, bullshiting, deception, lying, and concern with the truth of what one says. *Pragmatics & Cognition*, 23(1):53–67.

Clapp, L. (2020). Stalnaker on the essential effect of assertion. In Goldberg (2020).

Del Pinal, G. (2015). Dual Content Semantics, privative adjectives, and dynamic compositionality. *Semantics and Pragmatics*, 8(7):1–53.

Del Pinal, G. (2018). Meaning, modulation, and context. *Linguistics and Philosophy*, 41(2):165–207.

Dretske, F. (1988). *Explaining Behavior*. MIT Press, Cambridge, MA.

Fallis, D. (2024). Deceiving versus manipulating. *Analytic Philosophy*, 65(2):223–240.

Frankfurt, H. G. (1988). On bullshit. In *The Importance of What We Care About*. Cambridge University Press, Cambridge.

Goldberg, S. C., editor (2020). *The Oxford Handbook of Assertion*. Oxford University Press, New York.

Green, M. S. (2009). Speech acts, the handicap principle and the expression of psychological states. *Mind & Language*, 24(2):139–163.

Guerrini, J. (2024). Keeping *fake* simple. *Journal of Semantics*, 41(2):175–210.

Kotzee, B. (2020). Bullshit assertion. In Goldberg (2020).

Marsili, N. (2024). The definition of assertion. *Mind & Language*, 39(4):540–560.

Marsili, N. and Green, M. S. (2021). Assertion. *Journal of Pragmatics*, 181:17–28.

Pagin, P. (2011). Information and assertoric force. In Brown, J. and Cappelen, H., editors, *Assertion*. Oxford University Press, New York.

Pagin, P. (2020). The indicativity view. In Goldberg (2020).

Searle, J. R. (1986). Meaning, communication, and representation. In *Philosophical Grounds of Rationality*. Oxford University Press, New York.

Siebel, M. (2020). The belief view of assertion. In Goldberg (2020).

Sorensen, R. (2011). What lies behind misspeaking. *American Philosophical Quarterly*, 48(4):399–409.

Sperber, D. and Wilson, D. (2026). Rethinking ostensive communication in an evolutionary, comparative, and developmental perspective. *Psychological Review*, 133(2):315–338.

Stalnaker, R. C. (1999). Assertion. In *Context and Content*. Oxford University Press, New York.

Stalnaker, R. C. (2014). *Context*. Oxford University Press, New York.

Stokke, A. (2018). *Lying and Insincerity*. Oxford University Press, New York.

Wilson, D. and Sperber, D. (2012). *Meaning and Relevance*. Cambridge University Press, New York.

Wreen, M. (2013). A p.s. on b.s. *Metaphilosophy*, 44(1–2):105–115.

Wright, C. J. G. (2003). Realism, anti-realism, irrealism, quasi-realism. In *Saving the Difference*. Harvard University Press, Cambridge, MA.