

The Big Shill

Robert Mark Simpson and Eliot Michaelson¹

Ratio 33/4 (2020): 269-80

Abstract. Shills are people who endorse products and companies for pay, while pretending that their endorsements are ingenuous. Here we argue that there is something objectionable about shilling that is not reducible to its bad consequences, the lack of epistemic conscientiousness it often relies upon, or to the shill's insincerity. Indeed, we take it as a premise of our inquiry that shilling can sometimes be sincere, and that its wrongfulness is not mitigated by the shill's sincerity, in cases where the shill is sincere. Our proposal is that the shill's defining characteristic is their knowingly engaging in a kind of speech that obscures a certain aspect of its social status – most commonly, by pretending to speak on their own personal behalf, while in fact speaking as an employee – and that this sort of behaviour is objectionable irrespective of any other features of the shill's conduct. This sort of obfuscation undermines a socially beneficial communicative custom, in which we conscientiously mark the distinction between personal speech and speech-for-hire.

NOUN: *shill* | PLURAL NOUN: *shills*

1. *An accomplice of a confidence trickster or swindler, who poses as a genuine customer to entice or encourage others, e.g. 'I used to be a shill in a Reno gambling club'*

– OXFORD ENGLISH DICTIONARY

1. Introduction

We know a shill when we see one: the sales rep for a pharmaceutical company who tells doctors that a new opioid is only mildly addictive; the climate scientist whose research is sponsored by a large oil company; the social media influencer hawking 'Flat Tummy Tea'. To a rough first approximation: the shill is someone paid to speak in support of a product, organisation, or cause, while pretending to be speaking ingenuously, on their

¹ This work is entirely collaborative and names are given in reverse alphabetical order.

own behalf.² But what makes shilling objectionable? It often involves lying, or other forms of deception. Some shilling has bad results, and is thus objectionable on that basis. Probably most shilling involves a *prima facie* objectionable failure of integrity. The more interesting question is whether there is something about shilling as such that makes it *intrinsically* objectionable, over and above the garden variety wrongs that a particular shill commits in his particular context.

We take it that many people do intuitively regard shilling as wrong in this sense, and that they see its wrongness as being tied to the fact that the shill, unlike the candid spokesperson, obscures the nature of their relationship with the actor on whose behalf they are shilling. We want to defend this embryonic intuition. The problem with shilling, we will argue, is that in muddying the boundaries between ingenuous speech and speech-for-hire, the shill undermines a particular kind of salutary communicative social ethos that we should generally try to sustain.

Our interest in these questions is partly driven by the rise of influencers on social media platforms like Instagram. Not all influencers are shills. Certain influencers appear to be interested in fame for fame's sake. But many influencers resemble 'old-fashioned' corporate shills, in acting as paid spruikers who pretend that their spruiking is ingenuous and unpaid. Part of what makes our topic here intriguing is the difficulty we face in understanding where people's intense distaste for influencers is coming from. Yes, people find them obnoxious, and see them as a bellwether of some technologically-mediated cultural decline. But this isn't philosophically newsworthy. What would be newsworthy would be if influencers were communicating in a way that was distinctive, and distinctively wrongful. Initially it may seem like there is a complex form of insincerity involved in their trade. But, in fact, influencers can be sincere, and in some cases they are. The influencer's speech sometimes seems more objectionable precisely because of its sincerity. In giving an account of what shilling is, and of what makes it wrong, we want to get to grips with this peculiar subtlety.

In §2 we argue that shilling's objectionability cannot be reduced to a complaint about insincerity. In §3 we argue that its objectionability cannot be boiled down to a complaint about epistemic negligence. In §4 we elaborate and defend our central claim that shilling undermines a socially valuable communicative ethos, by blurring the boundaries between ingenuous speech and speech-for-hire. In §5 we explain why the same sort of charge shouldn't be levelled against spokespeople or innocent product enthusiasts, despite the fact that they commonly speak in ways that resemble the shill's speech, in certain respects. We conclude in §6 by discussing the facilitation of shilling on social media platforms, and the phenomenon – an ominous phenomenon, we want to argue – of aspiring influencers acting like shills for no charge, in the hope of attracting paymasters.

² This is how the term is typically used in ordinary parlance, today. Some dictionaries define the term in a slightly antiquated way, as referring to someone acting as a con artist's accomplice.

2. Shilling and Insincerity

To start with we want to examine the idea that shilling is just a (slightly complicated) way of being insincere, and that its objectionability owes to its insincerity. To set the ball rolling with this, it will be helpful to get a few cases in view.

BIG PHARMA. Molly is a drug rep for Big P Pharma.³ She is told by her employer that Opioid X is only minimally addictive. This is what she tells doctors at the conventions she attends, when she takes them for fancy dinners, etc. Conveniently, Molly also has a degree in pharmacology, so she is capable of reading the relevant research. She makes a point of telling the people that she's repping to that she is not just hawking Opioid X for pay: the data actually show it to be a safe, effective treatment.

Version 1: Molly has in fact read the relevant research and believes her employer's summary of it is inaccurate. She believes Opioid X is highly addictive.

Version 2: Molly has read the relevant research and she believes her employer's summary of it is accurate; while there is room for doubt, she believes that the preponderance of evidence suggests that Opioid X is only minimally addictive.

Version 3: Molly has access to the relevant research but hasn't bothered reading it properly. She takes her employer's summary of the known facts at face-value, and thus she accepts and attests that Opioid X is only minimally addictive.

Here is another structurally similar set of cases.

BIG OIL. Rex is a respected climate scientist. After building climate models for decades his work has taken an unexpected turn. Against the climatological consensus, and his own earlier work, Rex's models now indicate that the Earth's warming isn't driven by human activity. Rex's recent research is funded by grants from SpillHappy Oil, but he takes measures to obscure and downplay this fact as far as possible.

Version 1: Rex does not believe that his recent climate models are accurate. Rex's models attribute the Earth's atmospheric warming to increased solar flare activity, but he doesn't believe this explains the relevant explananda. Rather, he has been brazenly cynical in rigging his models to attract grants from SpillHappy Oil.

Version 2: Rex believes that his models are accurate. Rex's models attribute the Earth's atmospheric warming to increased solar flare activity, and Rex takes this to be a plausible account of things. Other climatologists have failed to persuade him that his models are essentially mistaking correlation for causation. Rex is confident that the others are in fact mistaken, and that he's one of a small number of climatologists whose beliefs on this matter tally with the facts.

Version 3: Rex accidentally encoded an exaggerated causal connection between solar flare activity and atmospheric warming into his models several years ago, and this has propagated through descendent models. If Rex were to thoroughly scrutinise his own

³ We originally used the names of real companies, but we were asked by the journal to change these. We leave it to the reader to run an appropriate mental 'find-and-replace'.

code, he would regard this as an error. But he is so convinced of his superiority as a climatologist that he does not feel disposed to do so.

We take it that, in each version of these cases, the shill does something wrong, something objectionable. We take it, further, that this would have been the case even if what they were saying actually turned out to be true. We will return to this later, in §4. Our more immediate interest is to point to what we take to be the primary take-away from these cases: given how the details vary across the different versions, it turns out that a shill's statements are not necessarily insincere. Obviously, a shill could say things that she believes to be false – as we see in version 1 of each example. But, equally, a shill can believe what she says. Even if what she is saying is false, or even demonstrably false, the shill could be credulous, or dogmatic, or apathetic about critically examining her beliefs. In short, then, shilling cannot be felicitously classified as a species of insincerity.

Perhaps this line of argument rushes to its conclusion too quickly though. Consider versions 2 and 3 of BIG PHARMA or BIG OIL. Although our shills have a surface-level belief in what they are saying, it also seems like they must, at the same time, harbor a deeper awareness of the possibility of their surface-level beliefs being badly-founded. In BIG PHARMA version 3, for example, Molly knows full well that she hasn't done her homework. Surely this awareness must register somehow or other in how she thinks about the addictiveness of Opioid X. Perhaps then there is a sense in which Molly is of mixed mind about what she believes.

This line of reasoning finds *prima facie* support in recent discussions of insincerity. On the orthodox, Searlean analysis of insincerity, an assertion is sincere iff the speaker believes the asserted content (see Searle, 1969, pp. 66-67). A number of challenges have been offered to this analysis, but relevant for our purposes is one due to Michael Ridge (2006). Ridge's key insight is that sincerity can co-exist with self-deception, in a way that a Searlean analysis overlooks. John might believe that he believes that his mother loves him, while, deep down, believing that she does not. Suppose that, when asked whether his mother loves him, John says 'yes, of course she does'. Intuitively, John speaks sincerely. But the orthodox account of sincerity regards his remark as insincere, because John has asserted that p despite his believing not-p (Ridge, 2006, pp. 488-89).

Andreas Stokke (2014) suggests that we can account for cases like this if we replace the orthodox *believe-what-you-assert* criterion with an alternative one. In order to sincerely assert that p, a speaker has to *mentally assent* to p's content. What is mental assent? According to Stokke, it means having a conscious belief-like attitude to the asserted thing. To illustrate, consider our earlier case: John consciously affirms that his mother loves him, even though, unconsciously, he believes the opposite. In other words, on Stokke's view, John's sincerity is explained by his having only one relevant conscious belief: namely, that his mother loves him (Stokke, 2014, pp. 504-08).

Suppose for the moment that Molly is similarly conflicted in version 3 of BIG PHARMA. Might she nonetheless count as insincere in asserting that Opioid X is only mildly addictive? If so, we could potentially hold on to a simple, insincerity-based account of shilling.

Unfortunately, the insights from Stokke and Ridge actually nudge us the other way – that is to say, back to our initial appraisal of the cases. Why? Because Molly’s sole relevant conscious belief is that Opioid X is only mildly addictive – a belief which is in line with what she asserts. So whatever is distinctive about shilling cannot be captured by construing shilling as a species of insincerity.⁴

3. Shilling and Epistemic Negligence

Here is another possibility: maybe shilling is objectionable because the shill offers advice, endorsements, etc., without exhibiting the kind of epistemic conscientiousness that they would need to in order to be justified in this. In the case of insincere shill, this lack of conscientiousness will be blatant. But even when the shill believes what they are saying, it seems like this is sometimes reliant upon the shill making a point of not conscientiously interrogating their beliefs. Perhaps the objectionability of shilling is tied to this?⁵

The problem with this line of analysis is that there doesn’t seem to be any form of epistemic negligence that is distinctive of the shill. In this section we will run through a series of possible explanations as to what the sincere shill’s epistemically negligent performance might consist in. The basic problem with these various explanations is that none appears to track the phenomenon in the right way. Either they don’t encompass enough of the cases that we intuitively recognise as instances of shilling, or else they encompass things that don’t intuitively strike us as instances of shilling.

Consider: maybe what is going on with the sincere shill is that they are saying what they believe (fine), but they believe what they believe for bad reasons (bad). This clearly won’t suffice, though. People assert things that they believe for bad reasons all the time without counting as shills: think of people who buy into medical quackery, or various pseudosciences. And moreover, it is at least possible for the shill to believe what they believe for good reasons. In BIG PHARMA version 2, even if Molly’s beliefs are false, it appears possible that she could be believing as she does in a rationally acceptable way.

Here is another possibility. Maybe what’s going on with the sincere shill is that they say what they believe (fine), and they believe what they believe for bad reasons (bad, but not distinctive), and – but – they *know* that they believe what they believe for bad reasons. This is more promising, but it still does not align with the intuitive conceptual scope of ‘shill’. For example, in the version 2 versions of our cases, Rex and Molly take themselves

⁴ We are assuming that bullshitting and insincerity are distinct (if sometimes overlapping) categories. If bullshit constitutes a species of insincerity, this would complicate the argument slightly. But likewise, it would complicate any attempt to explain the wrongness of insincere speech.

⁵ As far as the underlying concerns go, there is a resemblance between the complaint against shilling outlined here, and W. K. Clifford’s famous contention that it is wrong, always and everywhere, for anyone to believe anything on the basis of insufficient evidence; see Clifford (1877) in Madigan (Ed.) (1999). We are entertaining the notion that it is unethical to *advise* people on the basis of insufficient evidence, whereas Clifford is saying that it is unethical to *believe* on the basis of insufficient evidence. But the explanation of the ethical concern is similar in the two cases.

to believe what they believe for good reasons. They may be mistaken about this, but nevertheless, the fact that they believe themselves to believe what they believe for good reasons is incompatible with them knowing that they believe what they believe for bad reasons.⁶ Moreover, in the version 3 versions of the cases, Rex and Molly are, for reasons of epistemic apathy, somewhat oblivious as to whether they believe what they believe for good reasons. It seems incorrect to us to say that Rex and Molly know, in this situation, that they believe what they believe for bad reasons.

Perhaps this could be rectified by saying that sincere shills are *in a position to know* that they believe what they believe for bad reasons. Something in this suggestion is worth pausing on. There are some people who have no qualms about saying brazenly insincere things for money. But most of us are not like that. Most of us – if we found ourselves positioned, in life and the labour market, such that our one pathway to material security were to take up shilling – would hope that the things we ended up being paid to say were things that we could sincerely sign onto. But many of the things that shills are paid to say are things that an epistemically conscientious person – someone who wants to review the quality of their reasons for believing what they believe, when they are in a position to do so – will not be able to sincerely sign onto. So far as that is the case, the viability of shilling as a profession will depend upon people being epistemically unconscientious, in the way we have described above.

However, even this trait doesn't seem to carve at the concept's joints. Plenty of people who are epistemically unconscientious in this fashion don't engage in the shill's distinctive form of communicative disingenuousness. There are all sorts of people who have formed beliefs on the basis of bad or insufficient evidence, who are in a position to know that this is the case, and yet are not shills. Asserting various things that one believes while failing to engage in diligent doxastic-record keeping vis-à-vis the grounds of those beliefs is a completely run-of-the-mill failing in societies like ours, for beings like us with an ordinary suite of cognitive limitations.

These thwarted proposals might lead us to a broader worry: perhaps there is, in fact, nothing distinctive about shilling. Whether or not they are sincere, shills are in a deeply epistemically compromised situation. There are all sorts of things about their life and livelihood that impair their ability to have an ingenuous doxastic take on the matters on which they're paid to shill. However, so one might argue, this situation is not *that* different to everyone else's. All of us are subject to a range of factors – advertising, political ideology, pressures of social loyalty – that undermine our capacity to have an ingenuous doxastic take on matters. To use the terms favoured by some Marxists, false consciousness and ideology are likely to be the norm, rather than the exception, for beings like us in societies like ours.

There is at least a grain of truth in this. Most of us, possibly all of us, have at least some beliefs that are borne of us trying to rationalise our social position and perspective, in

⁶ This follows straightforwardly from the widely-accepted view that in order to know that p one must believe that p. A cannot *know* that they believe that p for bad reasons, unless they *believe that* they believe that p for bad reasons. And A cannot believe that they believe that p for bad reasons, while believing – as Rex and Molly both do, in version 2 – that they believe that p for good reasons.

ways that neglect to take account of various relevant explanatory factors that aren't readily apprehensible from a first-personal viewpoint. But even if we grant as much, there still appears to be something distinctive about the epistemic position one finds oneself in as a shill. This individual self-consciously participates in the distortive processes that colour her doxastic affairs. She is proactively complicit in these processes, in a way such that she will either be acutely aware of her having been epistemically compromised, or will have to enter into some kind of state of self-deception about this.⁷ What's more, the shill is in the relatively unusual position of being able to pinpoint most of the particular token beliefs in relation to which she has become epistemically compromised.⁸

The point of this is not to suggest that shills are distinguished by their asserting things that they've been epistemically negligent in forming beliefs about. Shills aren't necessarily epistemically negligent in the above ways, and in cases in which they are epistemically negligent, their negligence is often of a kind that people can exhibit without being shills. On the other hand, there is a form of epistemic misconduct that it seems like only a shill could be guilty of. Consider a case in which a person, A, is being paid to assert that p, while wanting others to believe that she believes p ingenuously, only for her to then come to believe that p, while also gulling herself into thinking that her coming to believe p was not causally related to her getting paid to assert that p. There is something almost grotesque in this, with A simultaneously playing the roles of both the mark *and* con artist. It is as if she is getting paid not just to mess with other people's attitudes vis-à-vis p, but her own too.⁹

Not all shills end up in this baroquely intellectually compromised state. Some simply engage in plain old lying. Others remain aware of the ways in which they have distorted their own epistemic position in taking up their trade. But again, it seems that this particular form of epistemic self-destruction is one into which only a shill could descend. And in reflecting on what makes it possible for the shill to descend into this state, we see what it is that makes shilling a distinctive, and distinctively worrisome, communicative practice.¹⁰

⁷ Granted, some authors will say that in fact most people in positions of social privilege are to some extent 'proactively complicit' in distortive processes that colour their beliefs – that they are, to use a term favoured by Medina (2012), *actively ignorant*. Even if that is right, though, there is presumably still something distinctive – something abnormally self-transparent, we might think – about the type of active ignorance that sincere shilling typically necessitates.

⁸ One might worry that such an ability to pinpoint these beliefs is incompatible with the possibility of self-deception regarding their epistemic status. We don't have space to argue for the contrary here, but see Jordan (forthcoming) for considerations that we take to cut in favour of this claim.

⁹ One might be tempted to say that shills of all stripes interfere with their ability to say what they believe because they believe it; on the importance of this relation, see Hannon (2019). While we are sympathetic to this line of thought, we are also unsure how to adjudicate such claims in the present type of case.

¹⁰ We are assuming here that it is possible for the shill to be *innocently* self-deceived, i.e. self-deceived in a way that wasn't consciously embraced at any point. But we grant that assumption might be mistaken. After all, the shill has to remain acutely aware of the strategic purposes that her communication is meant to fulfil, where those purposes are determined by her paymaster's aims.

4. The Distinctively Objectionable Thing about Shilling

Here is our proposal, then. Shilling isn't objectionable because it is insincere, or because it involves a shortfall of epistemic conscientiousness. Rather, the feature that defines shilling – speaking in support of a product, organisation, or cause, in a manner designed to obscure that fact – is objectionable in and of itself. Or to restate this using a few bespoke terms of art: it is objectionable to speak *clerically* – that is, representing a point of view other than one's own, for pay – while taking steps designed to undermine a listener's ability to classify one's speech as clerical, as opposed to *personal*. Prototypically, such obfuscation is carried out by *pretending* to speak personally while actually speaking clerically, but it seems to us that other methods of obfuscation will be possible as well.

To see why we might regard it as objectionable for someone to pass off clerical speech as personal speech, it will be useful to have an example that brackets off some of the confounding factors that might have coloured our intuitive judgements about the objectionability of shilling. To this end, consider the following case.

RENEWABLES. Zara is a sales rep for a renewable energy company, SolarBonanza!, which is tendering for major government contracts. SolarBonanza!'s products really are both less toxic to produce and more efficient than the alternatives. Zara has relevant education and expertise, and having reviewed the technical information and modelling on SolarBonanza!'s products and rival companies' products, she has a justified belief in the superiority of SolarBonanza!'s products. She makes a point of telling the government reps she meets that she isn't just hawking SolarBonanza! for pay; their products really are the best. But this is merely a tactic she employs to make her sales repping efforts more effective. In fact, for her part, Zara is a disillusioned, misanthropic anti-natalist. Her motive for repping for SolarBonanza! well is purely monetary. As far as the benefit to the environment goes, she feels something between irritation and indifference.

We take it that the expected consequences of Zara's shilling are overall positive: if her shilling succeeds, more renewable energy will be generated with fewer harmful side-effects. So shilling need not be harmful overall.

What's more, Zara believes the claims that she makes on SolarBonanza!'s behalf, and she has good reasons for believing them. So whatever intuitive objection we feel towards Zara's conduct, it cannot be based on insincerity, epistemic unconscientiousness, or an expectation of harmful consequences. It seems like it must be the brute fact of Zara's pretence – vis-a-vis the personal character of her speech – to which we are intuitively responding.¹¹

¹¹ Again, we do not mean to suggest that pretense is essential to shilling. Consider a variant on our example where Zara instead says: 'Listen, I'm paid to endorse SolarBonanza!. My bosses don't care whether SolarBonanza! is the best, or whether I believe it is, they just pay me to claim that it is. But I'm telling you: I've done my homework and SolarBonanza! really *is* the best. I get it, you're thinking "that's just what she *would* say, to try to persuade us". But would I be *levelling with you*, like I am now, if I wasn't really sincere? I'm paid to say this, but that actually has no bearing on why I believe it. I believe it because the evidence supports it.' This sort of *transparent shill* isn't misrepresenting her self-understanding. Rather, the problem is that she is failing to acknowledge that, from the listener's

Is our intuitive condemnation of this pretence justified, or is this merely a kind of customary prejudice, something like a response to a breach of etiquette? If it is justified, it is justified insofar as we have a reasonable expectation, subject to various caveats, that people will show us their true face, not only in saying what they believe – Zara fulfills that expectation – but in letting us see their real interests, attachments, and preferences, i.e. the ones they hold independently of their formal occupations and public roles. One important caveat to this generalisation is that no one is obliged to show her true face to any particular person who wants to see it. It is permissible – obviously, uncontroversially – to guard one’s privacy and time, and to relate to other people in a politely arms-length fashion. What is objectionable, however, is to tactically allow another person to believe that one is revealing one’s true face to them, when one is actually speaking clerically, i.e. delivering an advertisement or PR message on behalf of an organisation.¹²

One way of theoretically positioning this claim about the objectionability of shilling is to assimilate the shill’s pretence to other kinds of lying and misleading. In keeping with a fairly mainstream take on these matters, we could propose that the shill’s pretences, much like garden variety lying, wrongfully manipulate those to whom they’re conveyed. There is presumably something to this kind of analysis. The shill is indeed trying to manipulate his audience in some shape or form. And notice also that some garden variety liars and misleaders have a hard time recognising themselves as engaged in a manipulative endeavor. The shill may be more likely to consciously intend to perform a manipulative manoeuvre.¹³

This assimilating approach would ask us to give up the *prima facie* plausible idea that there is something distinctively wrong about shilling. Shilling remains distinguishable as a type of communicative activity, but our objections to it become generic. So we have at least a *pro tanto* reason to look for another way of theoretically positioning our explanation. Another unsatisfying thing about the assimilating approach is that it doesn’t readily account for the nebulous sense of personal affront that one might feel towards shilling that isn’t addressed to oneself. Even if you’re no part of Zara’s audience in a case like RENEWABLES, you may still feel a sense of indignation towards her conduct. Zara isn’t manipulating *you*, and thereby disrespecting *you*. And yet the indignation lingers. There is a sense that the shill – even a relatively benign shill, like Zara – isn’t just wronging her addressee, or committing an act which is wrongful from some impersonal moral point of view, but that she is somehow wronging everyone in her community. This intuition is difficult to make sense of if our objections to shilling are assimilated to our objections to other kinds of deceptive communication.

In trying to account for this, we will borrow an approach to thinking about communicative ethics, found in the work of Harry Frankfurt and Seana Shiffrin, which appeals to

perspective, what she is doing is indistinguishable from what a really clever, insincere shill might do to win her customer’s trust. Thus, we take it, she is still actively obscuring the clericality of her speech.

¹² In defending this idea, we are not endorsing a hard-line version of what Korsgaard (1986, p. 341) calls *the ideal of straightforwardness in human relationships*. All that we want to claim is that a certain form of guardedness in the service of a paymaster is objectionable – and no more than that.

¹³ For a recent, lucid account of this sort of mainstream view about the wrong of lying – in essence, that the wrong is about the liar unjustifiably manipulating the recipients of her lies – see Faulkner (2007).

a notion of corruption. Frankfurt (2005) famously argues that bullshit is importantly distinct from lying or deception. Unlike the liar, the bullshitter may believe that what he is saying is true. The issue is that the bullshitter doesn't care whether or not what he is saying is true. He says what he says because saying it conduces to some other purpose that he has, which is at best accidentally connected to his speaking truthfully. Frankfurt claims that the bullshitter is a greater threat to the values of truthfulness than the liar. The liar still recognises truth as a standard to which her speech is answerable, even if she is refusing to answer to it on this occasion. The bullshitter simply does not care. And this is a threat to truthfulness, for the same reason that a detached, ironic nihilism is a greater threat to public morality than brute wrongdoing. Our ideals can withstand violations, but they're shattered by indifference. Bullshit in public discourse has a corrupting effect on society's communicative practices, undermining, as it does, certain normative standards that help make those practices functionally useful (Frankfurt, 2005, pp. 59-61).

Shiffrin's accounts of lying and promissory infidelity also come with an idea of corruption working in the background. In trying to explain why we should sometimes keep promises made under duress, and why we should refrain from certain types of lying, even in 'murderer at the door'-type cases, Shiffrin is guided by the idea that our capacity to fully treat each other as moral agents depends upon us having a deep-seated allegiance to communicative authenticity. We must know others' inner states in order to fulfil our duties to others, and we need robust conventions of honesty and promissory fidelity in order to gain this knowledge. We should prioritise honesty and promissory fidelity, even at considerable cost, in cases where lying and promissory infidelity would not merely infringe against, but positively undermine, a social ethos of communicative authenticity. This kind of reasoning is evident throughout Shiffrin's account of these topics; see for example her analysis of 'murderer at the door' cases (2014, pp. 43-46).

The thing we want to borrow from these authors is their argumentative scheme, rather than any specific conjectures they advance about the consequences of lying or bullshitting. Both Frankfurt and Shiffrin appeal to an idealised picture of a communicatively well-ordered society – one in which everyone recognises that truth matters, or in which everyone adheres to a communicative ethos that conduces to authentic relationships. The forms of communicative wrongdoing they examine are diagnosed *as* wrongdoing because of the way that they disturb this order. The implied causal dynamic is, approximately, that of a convention being unravelled: by acting contrary to a certain regularity of conduct, the wrongdoer weakens the reason that everyone else has to comply with that regularity.¹⁴

Something similar can be said about the corrupting effect of the skill's blurring of personal and clerical speech. There are good reasons to have a communicative framework in which we conspicuously differentiate these two regions of expression. Many of us take up occupations in which we are called upon to speak on an organisation's behalf, and

¹⁴ Here we are just meaning to advert to the general, schematic understanding of conventions, widely-held since Lewis's (1969) seminal work on the topic, that they are those regularities of conduct for which our reasons for complying with them are fundamentally bound up in the fact that others also comply.

often we are instrumentally-minded in how we inhabit such roles. That is to say, an individual takes on a role for the sake of a goal that is outside of the role – colloquially: one works to live, rather than living to work. Being able to partition our identities in this way – colloquially: being able to wear and remove different ‘hats’, as a worker, and a member of a private club, and an untethered individual – enables us to partake of various benefits in a certain kind of public life, without relinquishing a sense of our own individuality. So far as the distinctions between personal speech and clerical speech are blurred, the benefits of this arrangement are compromised. One isn’t troubled by speaking in one’s capacity as a worker, so long as one is able to retain a private self apart from that role, and take off one’s worker’s hat. But if one is expected to toe the company line at all times – or, for members of a religious organisation or political party: if one is expected to stay ‘on-message’ 24/7 – then the costs of assuming a role in public life can become overwhelming. To enter public life is, *ipso facto*, to renounce one’s enjoyment of a cordoned-off, private self.¹⁵

We also have something at stake *qua* listeners, in these communicative conventions, not just *qua* speakers. Sometimes you want to know what the official company or party line is on a given issue. Other times you’re interested in what your interlocutor thinks about that matter once all their hats are removed. As individuals with cordoned-off, private selves, we want to be able to enter into meaningful relationships with other people with cordoned-off, private selves. This relational desideratum is utterly thwarted in a world in which our interests, attachments, and preferences are fully dictated to us by our public roles. And it isn’t enough just to allow space for private selfhood. We also need to have communicative practices via which we can reliably distinguish personal and clerical speech. If you are always suspicious that your interlocutors are parroting a company line (or a party-political talking point, or denominational orthodoxy), then you are cut-off from authentic personal relationships, almost as much as if you were living in the world where private lives are subsumed by public roles.

¹⁵ Nagel (1998) is one author who argues that we should try to uphold social practices that don’t necessitate this kind of trade-off, especially when it comes to people standing for public office.

5. Near-Shills: Spokespeople and Naïve Enthusiasts

The salutary communicative framework that we are gesturing towards need not be a costly thing to sustain. It is arguably already supported in the law, in most liberal democracies, by virtue of them adopting a differentiated approach, in free speech jurisprudence, to speech which conveys a personal opinion, and speech that is conveyed on behalf of a corporate entity or organisation.¹⁶ It might also necessitate limits being imposed on what employers can contractually demand of their employees, in relation to their communicative activities ‘out of hours’.¹⁷ Beyond a few *de jure* norms of these kinds, the maintenance of the sought-after communicative framework in societies like ours will largely rely upon the *de facto* custom of people indicating to others when they are speaking in a clerical capacity. People often do exactly this. They say things like: ‘if I was wearing my SolarBonanza! hat, I would say that...’. In plenty of cases the context of a communicative act is enough to indicate whether a speaker is speaking clerically or personally. Clarificatory remarks are only needed when the context is ambiguous.

Having explained the objectionability of shilling along these lines, we can now explain why other kinds of communicative behavior that partly resemble shilling aren’t objectionable in the way that shilling is. Consider the acts of a corporate spokesperson. We may object to the spokesperson who works on behalf of a bad organisation, insofar as she becomes a party to their ends. In this regard a spokesperson and a shill may be in the same boat. But once this consideration is bracketed off, there is little to object to in being a spokesperson. In RENEWABLES, if Zara had refrained from any subterfuge about her personal investment in SolarBonanza! – if she hadn’t used the pretence of talking ‘off the record’ as a sales tactic – there would have been nothing to object to. Spokespeople who forthrightly indicate that they are speaking as spokespeople do not damage the distinguishability of personal and clerical speech. Quite the opposite: they fortify the communicative framework in which the two are distinguishable, by clearly marking out the clerical, impersonal character of their communicative activities.

A trickier case is the rank and file employee who, having worked for an organisation for some time, experiences some kind of melding of personal identity and public role. The

¹⁶ The idea here is that our framework is supported if free speech jurisprudence offers maximal protection to paradigmatic cases of personal speech, e.g. sincere avowals of one’s beliefs, while only offering more qualified protections to advertising, direct marketing, and other commercial speech. This kind of arrangement exists in the US. Regulations on most kinds of personal speech are subject to strict scrutiny, such that they’re ultimately, nearly always, adjudged unconstitutional, whereas – under the doctrinal principles enshrined in the landmark case of *Central Hudson Gas & Electric Corp. v. Public Service Commission*, 447 U.S. 557 (1980) – regulations on commercial speech are more readily permitted. In short, deceptive and misleading commercial speech are simply unprotected from government regulation, while regulations on non-misleading commercial speech are permitted provided that they directly advance the legitimate government interests that putatively warrant the regulation. Granted, the sincere shill’s speech creates a tension in this kind of jurisprudential approach, because while its personality might formally entitle it to full-blooded protections, intuitively, we judge that its ultimate commercial purposes should make it liable to government regulation on basically the same terms as other commercial speech, especially if and when it is misleading. Our arguments in this paper would support the closing of any loopholes through which companies are able to take advantage of more robust protections by laundering their commercial speech through a shill’s personal speech.

¹⁷ Anderson (2017) is one recent work that’s concerned with this issue, and related issues about the limits of employer control over employee conduct ‘out of hours’.

interests, attachments, and preferences that belong to this individual's public role may become a proxy for her personal attitudes. Suppose Joy is a software engineer at Goggle. She may experience a mild but ultimately inexorable psychic pressure to desire that Goggle will annihilate PricklyPear in the battle for tech supremacy. And this may result in Joy communicating in ways that bear a notable resemblance to paid shilling. Suppose that Joy attends a software convention and presents an enthusiastic, impromptu speech to a group of investors about why Goggle's software is far superior PricklyPear's. Does Joy's spruiking in this case conflate the clerical and the personal, in an objectionable manner?

We don't believe so. It is in one's interests that one be able to maintain a private self who stands apart from one's public roles. But one may then choose to align one's private self with one's public role. Living to work, rather than working to live, isn't necessarily objectionable – if one's work is valuable, and if one autonomously chooses such a life. Someone like Joy, who has an interwoven public-private self, will, like Zara, tell others that she isn't just singing Goggle's praises for pay, that she truly believes in Goggle. But unlike Zara, Joy isn't engaged in a pretence. She isn't disguising her true face. Rather, she's authentically manifesting to others the fact that – for better or worse – her personal face has merged with its clerical mask. Zara, by comparison, is duplicitously pretending that this is the case. And in this, she chips away at the normative force of the customary expectation that we make our personal and public selves distinguishable.

Another tricky case – though again, one that we see as importantly distinct from shilling, in terms of its ethical character – is the earnest, naïve enthusiast, who adores a product or company so much that he engages in shill-like spruiking on its behalf, without receiving any payment. Consider the following example.

GEARHEAD. Chris upgrades his crankset from an Ultegra to a DuraAce, shaving a cool 50 grams from his bike weight in the process. Chris takes several rides on his upgraded rig and boy does he feel faster! So much so that he takes to Facebook and announces: 'New crankset bros! WAY faster thanks to my new DuraAce.'

Chris vaguely resembles a certain kind of sincere shill: one whose belief in what they are paid to say is only sustainable thanks to their unconscientious disinclination to examine the grounds of those beliefs. If Chris were to thoughtfully reflect on the matter he would remember that his own body weight fluctuates within a 3-4 pound (1360-1814 gram) range, which effectively swamps the upgrade's effect on his bike's weight. In Chris's perception of his bike's performance, he is experiencing something like a placebo effect, and he would realise this if he reflected on things. Like the sincere shill, Chris is motivated to remain unreflective. But whereas the sincere shill's motivations are about avoiding cognitive dissonance in her occupation, Chris's motivations are about needing to rationalise the eye-watering sum that he parted with in acquiring his new crankset.

Again though, like Joy from Goggle, Chris is not disguising clerical attitudes as personal attitudes. If there is anything awry with him, it owes to the fact that his ingenuous personal attitudes are a bit more clerical in content than we feel they should be. This doesn't undercut the communicative customs under which we expect people to make their cler-

ical and personal attitudes distinguishable. Although perhaps, if the world became overrun with Chrises (or Joys), our reasons for wanting to uphold those communicative customs would be weakened.

6. Conclusion: Influencers and Wannabe Shills

We have argued that the intrinsically objectionable thing about shilling isn't that it is insincere, or epistemically unconscientious, but rather that it involves a pernicious obfuscation: the shill is speaking clerically, but intentionally takes steps to obscure that fact, typically (though not always) via pretence. Such obfuscation is objectionable because it undermines our conventional commitment to a particular form of communicative ingenuousness, one which helps mitigate the costs of taking on public offices and identities. The shill's pretences may be subtly self-deceptive, or more self-consciously cynical, but in either case the objectionable feature we have highlighted will still obtain. The *most* objectionable feature of a particular shill's activity might be the awful consequences that it contributes to, or the gross deception that it involves. But all shills, even those that don't perpetrate these particular wrongs, are liable to censure.

Many influencers on social media sites like Instagram are shills. They sell companies access to their feeds, offering some number of mentions or cameos with their product for a given fee. Influencers are rarely candid about these arrangements. (The efficacy of hiring influencers, as a marketing strategy, is, we surmise, inversely correlated with the influencer's transparency. To the extent that some influencer *is* candid about such arrangements, they would evade the charge of shilling.) We can take a well-known celebrity as an example. Khloe Kardashian regularly posts about a product called Flat Tummy Tea, a drink containing diuretics and laxatives, marketed to young women. In a typical post, Khloe will tell her followers: 'This @flattummytea is really working the bloat out!' We regard this as a paradigm case of shilling: Kardashian pretends to be praising Flat Tummy Tea just because. She avoids mentioning that she's a paid, undercover spokesperson.¹⁸ Presumably many or most of Kardashian's followers put two-and-two together, regarding her relationship with the company. But Kardashian keeps up the pretence all the same, because it is integral to the marketing strategy. And that pretence, we have argued, is one species of the sort of obfuscation characteristic of shilling.

Social media facilitates a potent shilling dynamic. People identify with celebrities as though they were friends, and social media provides a communication channel that aids that illusion, allowing the celebrity to retain an air of intimacy while hawking wares to a massive audience. Shilling may thus be in a process of becoming a stronger and more

¹⁸ More recently, many Instagram influencers have started listing the hashtag #Ad in their comments on their sponsored posts, though typically after numerous other comments and so well below the 'fold' and thus easily miss-able. Does this suffice to make these influencers something other than shills? We think not, as we take it that the fact that they are posting on what is otherwise their personal feed is more significant than this (often hidden) indication that a particular post has been paid for.

widely utilised marketing practice. Many people are instinctively troubled by that prospect, and the analysis put forward here about what makes shilling objectionable lends some measure of support to that concern.

One interesting phenomenon which is coming along for the ride with the expansion of social media influencing is what we would like to call *wannabe shilling*: people using social media accounts to post about products or companies, with the hope of getting hired as paid influencers, and/or scoring free schwag.^{19,20} We believe that wannabe shilling is a cause for concern, particularly in light of our account of what makes shilling objectionable. Thanks to a widespread perception of lucrative potential in becoming a paid influencer, it may be possible to simulate the negative effects of a shill's obfuscations without a formal shilling arrangement. The more moralistic reaction to this phenomenon would be to censure wannabe shills for chasing opportunities to take up an objectionable line of work. The more politically-minded reaction would be to train our critical gaze on the social forces – expanding wealth divides, a dwindling of meaningful, living-wage jobs, manipulative platform design by social media companies – that plausibly incentivise wannabe shilling.

Either way, the emergence of wannabe shilling doesn't augur well. It is despairingly easy to imagine a future in which the distinction between people's personal and clerical attitudes is viewed as a quaint relic. 'Remember the good old days, when you weren't contractually obligated to make posts on social media praising your employer's products and board of directors? Remember when we used to think of ourselves as having interests, attachments, and preferences that were in any sense independent of our paymasters?' Some critically-minded readers may balk at how this paper's analysis of shilling presumes that clerical and personal attitudes are in fact distinguishable. They may suggest that in societies like ours, where so much power lies in the hands of corporate titans, the untethered personal self is just a mythic construction that we use to delude ourselves about the provenance of our attachments and desires. They might claim that our attitudes are always already clerical. But this too would be an overstatement. The clericality of the individual's personality is not some all-or-nothing matter. An individual persona which is not dictated by one's public roles might be more or less sustainable at different historical moments, under different economic, cultural, and technological conditions. No rose-tinted view of present conditions is required in order to dread the conditions superseding them.²¹

¹⁹ Schwag (n.): Items given away free, typically for promotional purposes, to people attending an event, using a service, etc. (Oxford English Dictionary).

²⁰ In a somewhat tongue-in-cheek spirit, a philosopher in the UK has recently launched an Instagram account focusing on the variety of delicious-looking ramen bowls that he can prepare in his office using only Tesco house-brand ingredients. Each of his posts is tagged with '#tesco', in an attempt to garner Tesco's attention and sponsorship. Is this person a shill for Tesco? Not quite (or yet), because he isn't engaged in the relevant kind of obfuscation vis-à-vis personal v. clerical communication.

²¹ For helpful discussion and feedback, thanks to Herman Cappelen, Amanda Greene, Michael Hannon, Dan Harris, Torfinn Huvnes, Rob Hughes, Brian Hutler, Julian Jonker, Jessica Keiser, Maggie O'Brien, Jessica Pepp, Joey Polluck, Matthew Shields, Dan Singer, Rachel Sterken, and Elmar Unnsteinsson. Thanks also to our audiences at the University of Oslo, the Wharton School of Business, and the Philosophy Society at King's College London.

References

- Anderson, E. (2017). *Private Government: How Employers Rule Our Lives (and Why We Don't Talk about It)*. Princeton: Princeton University Press.
- Clifford, W. K. (1999). The ethics of belief. In T. Madigan (Ed.), *The Ethics of Belief and Other Essays* (pp. 70-96). Amherst: Prometheus. [Originally published 1877.]
- Faulkner, P. (2007). What is wrong with lying? *Philosophy and Phenomenological Research*, 75(3), 535-557.
- Frankfurt, H. (2005). *On Bullshit*. Princeton: Princeton University Press.
- Hannon, M. (2019). *What's the Point of Knowledge?* Oxford: Oxford University Press.
- Jordan, M. (forthcoming). Literal self-deception. *Analysis*.
- Korsgaard, C. (1986). The right to lie: on dealing with evil. *Philosophy & Public Affairs*, 15(4), 325-349.
- Lewis, D. (1969). *Convention*. Cambridge Massachusetts: Harvard University Press.
- Medina, J. (2012). *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and the Social Imagination*. Oxford: Oxford University Press.
- Nagel, T. (1998). Concealment and exposure. *Philosophy & Public Affairs*, 27(1), 3-30
- Ridge, M. (2006). Sincerity and expressivism. *Philosophical Studies*, 131(2), 487-510.
- Searle, J. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Oxford: Oxford University Press.
- Shiffrin, S. V. (2014). *Speech Matters: On Lying, Morality, and the Law*. Princeton: Princeton University Press.
- Stokke, A. (2014). Insincerity. *Noûs*, 48(3), 496-520.