

ORIGINAL ARTICLE

A moral critique of psychological debunking

Nicholas Smyth 

Department of Philosophy, Fordham University, New York, NY, USA

Correspondence: Nicholas Smyth, Department of Philosophy, Fordham University, New York, NY, USA.

Email: nsmyth1@fordham.edu

Posing questions with a *hammer* and, perhaps, hearing in reply that famous hollow sound that indicates bloated intestines... what a pleasure for me, an old psychologist and pied piper; in my presence, the very things that want to keep quiet are *made to speak out*.—

Nietzsche, “How To Philosophize With a Hammer”

I suppose it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail.—

Abraham Maslow

Suppose I am sincerely advising a friend on how to pursue her career in our organization. I encourage her to be mild, kind, and to seem unambitious around her superiors. A third party overhears our conversation and says to me: “You’re only saying those things because you’re afraid she’ll get the positions *you* want.” This is a paradigmatic case of a *psychological debunking*: a speech-act which expresses the proposition that a person’s beliefs, intentions, or utterances are caused by hidden and suspect psychological forces.¹

On the face of it, it can seem as if this third-party has done both my friend and I a service by offering a psychological debunking of my utterances. But has he? In this paper, I wish to shed some light on some generally unacknowledged dangers involved in this activity. Of course, psychological debunking is not intrinsically suspect or confused; the origins of our mental states and activities matter to us, and they should matter to us. However, the activity of debunking is hazardous from a moral and social point of view, and this is something that ought to concern us far more than it does. A great deal of work has been done on the conceptual and normative status of debunking arguments (Kahane 2011; Korman 2019; Mason 2010; Nichols 2014; Sauer 2018; Vavova 2015). Nevertheless, we lack a plausible and fully-worked out *moral* critique of this popular activity, and that is precisely what I aim to supply here.

Before proceeding, an important caveat: I am not here proposing that there be any legal or institutionally enforced sanctions on debunking speech. The question in this paper concerns the morality of debunking, or whether debunkers are blameworthy for their behaviour. Moreover, I will make no complaint whatsoever about individuals inquiring into their own problematic psychological histories. My target here is the interpersonal delivery of a psychological debunking, normally from one agent to another.

1 | THE NATURE OF PSYCHOLOGICAL DEBUNKING

Let me outline this target with a little more precision. Socio-historical debunking arguments aim to uncover untrustworthy or reprehensible cultural histories, and evolutionary debunking arguments reach back even further, arguing that some belief or value is undermined in virtue of its being influenced by Darwinian forces. When I use the term *debunking* here, I mean to refer only to a third style of debunking which claims that some belief or intention or utterance is the product of hidden and unsavory *psychological* causes.² It is not always easy to draw clear lines between these various types of explanation—after all, social causes will almost invariably run through human psychology, and evolutionary explanations often invoke irreducibly social concepts.³ However, in this paper I will focus on explicitly psychological debunking.

Moreover, a debunking of this sort does not *just* say that an utterance or action has a bad or suspect psychological cause. For example, to respond to the use of a racial slur by claiming that the speaker is motivated by racism is not to “debunk” the utterance, since because the utterance wears its eminently blameworthy motivations on its sleeve. This is just moral criticism, and it would be odd to write a paper opposing moral criticism as such. Debunking does more than simply name a suspect cause: it suggests that the suspect motivations behind an utterance or belief are *not* what they seem to be. Most utterances literally express certain mental states, in the sense that hearers are conventionally licensed to infer that the speaker has those mental states on the basis of what they said. A debunking suggests that the genuine motivations for an utterance or action are not what they literally express.

Thus, to return to the example given at the outset of the paper, when I say “you should try to seem mild, kind and unambitious if you want to succeed here,” I literally express both (1) the belief that this is actually a good strategy for my friend to pursue, and (2) an underlying motive of care or concern. The debunker, on the other hand, suggests that my utterance wasn't caused by any of this, rather, it was caused by my own egoistic selfishness and some devious strategic reasoning: “if I tell her this, *then* she won't get the positions I want.” Debunking essentially belongs to the “Hermeneutics of Suspicion,” a tradition famously christened by Paul Ricoeur, and it seeks, in the Freudian jargon, to go beyond manifest content in order to reveal latent content (Freud 1920; Ricoeur 2008).⁴

But the hermeneutics of suspicion is not confined to any particular discipline or social milieu. It is easily recognizable in everyday life and conversation. Moreover, archeological suspicion has even become modestly popular even in analytic moral philosophy itself.⁵ That being said, in this paper, I won't criticize high-level philosophical debunking arguments, because they are distinctly impersonal: they do not attribute bad or distasteful motives to particular persons, and they tend to be directed at philosophical positions. By contrast, the sort of psychological debunking I take issue with here is *targeted* at a particular person or persons.

Now, Nietzsche is of course the patron saint of psychological debunking. Christians, utilitarians, Kant, Schopenhauer, Wagner and Socrates are all dented by the Nietzschean hammer, which psychologizes nearly everything it opposes (Nietzsche 1887). In general, the sorts of causes he names in his explanations—resentful motives, wishful thinking, physiological, and psychological weakness, shame at physical ugliness—are almost never vindicatory; if the target of the debunking were to accept the explanation, he or she would almost certainly be thrown into humiliation and despair.

Some philosophers have learned to respond to Nietzschean debunking by citing either the genetic or the *ad hominem* fallacy, implying that such causal histories cannot show that any of our beliefs can be false. However, this response rests on a misunderstanding of what debunking is (Boudry et al., 2015; Joyce 2007). Debunking never aims to undercut the *truth* of anyone's beliefs, rather, inasmuch as it attacks cognitive states at all, it merely threatens to show that we lack justification for holding them and/or that the belief is produced by dangerous or undesirable motives. For example, my

assertion “meekness and humility will take you far in this company” *could* still be true, but if I am being driven to say it by egoistic or selfish motives, my friend has no reason to believe it, and they furthermore have reason to worry about my motives. No logical “fallacy” is committed because no conclusion about truth-values is even on the table.

However, it is illuminating to return to the text in which the phrase “genetic fallacy” first occurred. This is Morris Cohen and Ernest Nagel’s *An Introduction to Logic and Scientific Method*. There, the authors distinguish sharply between logic and psychology, or between “the science of valid inference” and “the systematic study of how the mind works.”⁶ They go on to complain about the argumentative strategy which involves an “*argumentum ad hominem*, a popular device to deny the logical force of an argument (and thus to seem to prove the opposite).” They suggest that “any argument whatsoever can be ‘refuted’ in this way by inventing some unfavorable psychogenetic account of how or why the proponent of the argument came to hold that view.” (Cohen and Nagel 1934, iv).

This text is one of the first outspoken attempts in the philosophical literature to resist Nietzschean-Freudian debunking stories, which were proliferating in both popular culture and in the academy. Cohen and Nagel were clearly quite frustrated with this development:

Thus the fact that a man is rich or poor, married or single, old or young, is frequently used as an argument to disprove the truth of the proposition he affirms, or to lend force to its contradictory. This has received a great impetus in recent times from popular psychoanalysis. (Cohen and Nagel 1934, 380)

As I have just mentioned, these authors were mistaken to think that psychological debunking aims to “disprove the truth” of anything, indeed, both Nietzsche and Freud were quite clear on that point.⁷ However, as I will now show, it is *not* a mistake to be wary (or, indeed weary) of the activity itself. While of course many debunkings may be perfectly reasonable and justified, we should also understand that psychological debunking is intrinsically linked to shame and humiliation, to social fragmentation and to the erosion of trust. We who have learned to enthusiastically play Nietzsche’s game badly need a sustained ethical critique of this activity.

Before launching into my own critique, however, I’ll review two prominent attempts at such a critique and explain why they are inadequate. My aim here is not to provide some kind of literature review. Rather, the two attempts I will discuss are important for my purposes because they share a problematic feature. For each of these accounts, it is causal explanation *as such* that is worrisome: the activity of explaining a person’s utterances or beliefs is supposed to be intrinsically problematic all on its own. As we will see, this cannot be right, since such explanations are ubiquitous and a core part of moral practice. So, after rejecting these diagnoses, I’ll go on to say why psychological debunking is actually problematic. My own account, unlike these accounts, will preserve the important truth that only *some* types of psychological explanation are morally hazardous.

2 | MACINTYRE AND RINI ON DEBUNKING

Alasdair MacIntyre, in *After Virtue* and elsewhere, suggested that the activity of what he called *unmasking* was bound up with a more general cultural shift toward emotive ethics, a moral system within which value-judgments are nothing more than mere expressions of preference, desire, or taste. For MacIntyre, emotive morality can only become dominant in culture which has become morally lost, or completely unable to find firm rational foundations for its values. The popularity of unmasking, for him, was a symptom of this larger sickness, along with the rise of indignation and rage as

characteristic moral emotions. As he wrote: “unmasking the unacknowledged motives of arbitrary will and desire which sustain the moral masks of modernity is itself one of the most characteristically modern of activities” (MacIntyre 1981/2013, 72).

However, since MacIntyre wrote, a wide array of philosophers have questioned the unargued premise at the heart of his critique, namely, that emotions such as anger or indignation can *never* serve as rational foundations for our values (D’Arms and Jacobson 2000; Nussbaum 2003; Tappolet and Rossi 2015). Moreover, it is just not at all clear that unmasking is essentially tied to the nature of moral judgment; a rationalist culture where moral judgments are the result of calm, rational reflection must also be constantly on guard against what they will call “irrational” influences; Plato, an arch-rationalist inhabiting what is allegedly a pre-emotivist culture, built an entire moral philosophy around the imperative to guard against such influences.

I conclude that debunking is not morally problematic because it is a symptom of some runaway emotivist malady within modern culture. For a more plausible idea, I turn to Regina Rini’s recent critique of psychological debunking. Rini claims that moral discourse itself is nothing short of a solution to a Kantian antinomy between two propositions:

Thesis: My moral judgments sometimes have the phenomenal character of necessity—as if I simply must do what morality requires of me.

Antithesis: For any particular moral judgment, there is some contingent causal history that fully explains why I feel bound to abide by it.

(Rini 2020, 373–4)

Moral agents, Rini claims, operate under the (tacit) recognition that both of these claims are true. This seems basically right. To return to the paper’s controlling example, when I advise meekness and humility to my friend, I feel driven to do so by a sense of urgency or perhaps necessity: *I must tell her this now*. It simply seems like the right thing to do. But then, the debunker informs me that this sense has a causal history. So, I must now confront the metaphysical contingency of this sense of urgency. Rini thinks that I am being forced to confront a severe tension involved in entertaining these two sorts of claims. What is that tension, exactly?

Rini insists that the two thoughts—practical necessity and the metaphysical contingency of that sense of necessity—are difficult to *think* at the same time, such that believing one tends to undermine the other. Interpersonally, this tension is said to show up in the famous Strawsonian distinction between participant and objective stances toward persons, and Rini’s moral critique of psychological debunking comes down to this claim: it is often disrespectful or boorish to inflict the objective attitude on others because it tends to undermine the sense of practical necessity that accompanies their moral beliefs. In Rini’s words, this.

amounts to selectively suspending external validation of their experience of phenomenal necessity. It exposes one person to a form of existential vexation that none of us can truly weather. For that reason, it amounts to a type of defection or disloyalty. A person who sincerely attempts to convey the experience that her will is necessarily bound by moral requirements will naturally feel oddly assailed when met with the suggestion that her values can be explained away as mere contingencies.

(Rini 2020, 384)

But this cannot be right, because it cannot distinguish between causal debunking and causal vindication. In the example given at the start of this paper, what matters is not *that* my sense of moral urgency has a causal history, but that it has a *problematic* causal history which it might have lacked. Indeed, I will almost certainly already think of my utterance of having a causal history: I will think of myself as advising meekness out of genuine care. But the debunker alleges the opposite, and that is why I feel “assailed.” So, it must be a mistake to assume that it is the causal history itself that constitutes the debunking, since I am already operating under a certain conception of what motivates me.

Here is another way to illustrate the point. Suppose the would-be debunker in my example says something quite different to our new colleague: “You should listen to this advice. I know this person, and what they say is coming from a genuine sense of care and from many long years of experience with this company.” According to Rini, since it highlights a “merely contingent” causal history, this assertion must *debunk* my initial assertion, and I must feel “oddly assailed” by it. Yet, quite the opposite is true: I will probably feel that this person is actually *praising* me, or kindly reminding me of my own compassion and intelligence. This is a kind of vindication of my moral agency, yet it is, in Rini’s terms, a “mere contingency” which, on her view, can only produce existential alienation. This simply doesn’t seem right.

This point carries over to any debunking attempt whatsoever, even one that targets our more basic values. In this example I value compassion and empathy, and if a ‘debunker’ tries to undermine my confidence in these values by informing me that I got them from my compassionate and empathetic parents who always made sure my needs were met, then I am hardly thrown into existential anxiety over this. Not all explanation is “explaining away”; if it were, the very idea of moral reasons-responsiveness (in the causal sense) would be incoherent.⁸

Rather than follow this unpromising route, we should acknowledge that *some* psychological causes seem troubling for a very simple reason: they, unlike other causes, reflect badly on our character and our moral reliability. If someone tells me that in their view I am *really* advising meekness for self-interested reasons, then I am suffering a form of *humiliation*: a loss of social standing with respect to some shared normative standard (Gilbert 1997; Miller 1995; Statman 2000). For almost all of us, the resulting cocktail of shame and loss of self-esteem can be deeply painful. Moreover, I am being given evidence that I am the sort of person who is generally morally unreliable; someone who will believe and act on motives that don’t track the moral truth, and who may not even consciously recognize or understand that this is what they are like.

This, I submit, is the surprisingly simple answer to our question, one that requires no detour through anti-emotivist metaethics or Kantian antinomies. Interpersonal psychological debunking is usually morally problematic because it is a kind of shaming or *verbal humiliation*. It is painful to think that we are bad people or unreliable moralizers, or even just to realize that others think this way of us. This sense is never generated by *mere* contingency, rather, a debunking points to a particular set of contingent causes which are troublesome or problematic, by our own lights or by the lights of the people we regularly interact with. Our powerful desire to be better than this explains why we are wounded by attempts at psychological debunking.⁹

Now, I want to register an important caveat. It is undeniably true that we don’t always experience pain or humiliation when our utterances or beliefs are debunked, even if we accept the causal story itself. This is normally because we do not share the evaluative standards of the person launching the debunking; we don’t think that the causal diagnosis actually reflects particularly badly on us. However, even when this is so, almost all of us share an aversion to being hated or denigrated by others. In expressing their own displeasure, the debunker may simply aim to humiliate us in this more basic sense. Indeed, this form of humiliation acquires more and more power as the debunking audience grows, something that should be clear to anyone who has paid any attention to the new and historically

unprecedented phenomenon of internet shaming (Ronson 2016). So, it is fairly clear that a debunking speech act is always an *attempt* to humiliate, even if the humiliation is not always successful.

We can now begin to see why psychological debunking is often a morally hazardous activity. Virtually any model of ethics is going to contain presumptive duties of non-maleficence, a duty to refrain from humiliating others unless there are good reasons to do so (Rachels 1997). This, I claim, establishes a moderate presumption against debunking. However, this does not mean that psychological debunking is *always* wrong, since a presumption like this can easily be defeated, and in what follows I will try to outline some conditions that must to be met if this moral reason is to be defeated.

3 | THE MORAL HAZARDS OF DEBUNKING

The first moral presumption against debunking is grounded in its normally being an act of humiliation. After all, as I've just suggested, in the case of a debunking one is choosing to *undermine* a person, either directly to the person themselves or to an audience of others. This is just what our would-be debunker in my initial example does when he suggests that jealous ambition is causing me to recommend humility to others. I argue that in choosing to do this, he must make four crucial assumptions:

1. (Epistemic) He is in fact *reasonably* sure of the real causal history behind my expressions and of the evaluation he is making of that history,
2. (Evaluative) The end-state at which he is aiming (in this case, the exposure of hypocrisy) justifies his decision to undermine a person in this way,
3. (Means-Ends) Psychological debunking is actually an effective way to secure this valuable end given the context, and
4. (Unintended Consequences) His ends, even when efficiently achieved, are valuable enough to justify the potentially negative consequences produced by his chosen means.

If any of these conditions are not met, then our debunker is acting in a vicious or reprehensible manner. I'll now take each of these hazards in turn, showing how easy it is to make these morally serious errors.

3.1 | Epistemic hazards

Debunking, as we have seen, is an intrinsically speculative activity: one is claiming to know what is *really* behind some utterance or action, and this rather obviously raises the question of how it is that we can know such a thing about another person. In general, it is easy to get the impression that a huge amount of psychological debunking fails to meet even undemanding epistemic or interpretive standards. Very often, we simply do not feel the need to even collect evidence for the moral and psychological hypotheses which lie at the heart of our favored debunking stories. Rather, we simply share them in the company of others who share our sense of the obviousness of the causal story, and who will reward us with that knowing nod (or perhaps that "like" on social media). This results in the situation that Annette Baier eloquently described when she wrote that debunking "is a game that has become boring due to the laxness of its rules... it is easy to turn it on anyone and everyone" (Baier 1985, 231).

This, recall, was roughly one of the concerns motivating Cohen and Nagel: even as of 1934, debunking had become *too easy*. And it is hard not to see the sense in this point: even with our huge

array of well-developed social-psychological tools, we rarely feel the need to inquire into our grounds for even believing in some disreputable causal history, let alone in offering one in the public domain. Since this form of scientific inquiry is difficult and can take a great deal of time, and since our need to attack our opponents is so urgent, most people simply offer the disreputable causal histories and go about their business. This is seriously problematic from several normative perspectives.

To begin with a broadly consequentialist perspective, the general acceptance of low epistemic standards for debunking has had (and will continue to have) disastrous consequences for civil society and for untold millions of ordinary people. Seduced by the apparent obviousness of our own favored debunking stories, we forget that various other segments of society may rely merely on their own sense of obviousness to prop up devastating attacks on people and on institutions. As of this writing, it appears as though nearly a quarter of U.S. residents believe the frankly ludicrous claim that the 2020 election was “stolen,” and that claims to the contrary on the part of major figures are motivated purely by the desire to seize power. A segment of the global population believes that the dangers of the COVID-19 virus have been vastly inflated by key figures whose true aim is to establish Orwellian control over the general populace. Parents of children murdered in schools by gunmen are regularly forced to confront “skeptics” who claim that their children are still alive and that the parents’ claims to the contrary are merely motivated by the financial rewards that accompany collusion with a conspiracy. And, of course, we regularly hear it suggested that climate change is a hoax (somehow) perpetrated by thousands of scientists in a bid to secure power and influence.

None of these stories is intrinsically flawed; in each case the all-too-human motivations that feature in them are real, and certainly *could* explain what they purport to explain. But there is simply no direct evidence that this is so, and in each case this evidentiary gap is filled, psychologically, with that problematic sense of obviousness. In short, people die, people are subjected to horrendous abuse, people have their civil institutions dangerously weakened, and speculative psychological debunking is a precondition for all of it. After all, these agents must explain how so many apparent experts deny what they confidently assert, and references to shadowy motives are easily available. While we should have enormous respect for such thinkers as Nietzsche, Freud, and Foucault, we should also recognize that these conspiracy theorists are their bastard offspring, gleeful practitioners of the hermeneutics of suspicion.¹⁰ As a rueful Bruno Latour wrote:

Of course, we in the academy like to use more elevated causes—society, discourse, knowledge-slash-power, fields of forces, empires, capitalism—while conspiracists like to portray a miserable bunch of greedy people with dark intents, but I find something troublingly similar in the structure of the explanation, in the first movement of disbelief and, then, in the wheeling of causal explanations coming out of the deep dark below. What if explanations resorting automatically to power, society, and discourse had outlived their usefulness and deteriorated to the point of now feeding the most gullible sort of critique?

(Latour 2004, 229)

Thus, while our society desperately needs these “gullible” people to confront such things as publicly available facts and testimony from first-hand witnesses, they choose to dismiss such facts and perspectives *a priori* on the basis of speculative and ready-to-hand psychological debunking, a tool which our culture has rendered all too available.

So, there are powerful consequentialist reasons for trying to raise the epistemic bar here. Moreover, recent work on assertion suggests that psychological debunking, qua moralized assertion, may well be *constitutively* subject to certain epistemic norms, irrespective of any consequences produced by

the speech itself. The precise nature of these norms is disputed; Timothy Williamson has influentially argued that all assertions are subject to a *knowledge norm*, such that speakers ought only to assert what they know to be true. To this, Mona Simion has added that *moral* assertions must meet even more demanding norms; in addition to the knowledge norm, moral assertion must contain an explanation for *why* its constitutive claim is true. She claims:

At a context C, one's moral assertion that p is epistemically permissible only if (1) one knows that P and (2) one's assertion is accompanied by a C-appropriate explanation why P.

(Simion 2018, 483)

This is because, she says, the constitutive aim of moral assertion is to produce moral understanding, and to understand P is, in part, to know *why* P is true. Max Lewis has rightly argued that Simion's condition is too strong: what is needed is not an explicit *articulation* of the relevant moral explanation, but something in the conversational context that renders the explanation salient (Lewis 2019). In the vast majority of cases, the context allows hearers to get a grip on the intended normative explanation: if you tell me that meat-eating is wrong while we are watching a video on cruelty to animals, you don't need to say much at all to render your intended explanation salient.

Now, I think we should tread carefully here, because there is room for disagreement over the strength of these norms and of just what they require. You might think that, for example, a person need only transmit partial understanding of their evaluative point, and that they need only possess *some* epistemic justification for their causal and normative claims, even if this does not rise to the level required to count as knowledge. Moreover, you might think that such norms only apply to assertions which *matter* in some sense, and that relatively trivial assertions are not governed by such norms. Nonetheless, it does seem clear that in the case of debunking assertion, a person violates some set of important norms when they launch the assertion either (1) in the absence of evidence for the causal hypothesis, or (2) in the absence of any shared understanding for why the psychological causes named in the assertion are ethically suspect. Let's call this the *Simion-Lewis Point*, and it establishes basically deontic restrictions on debunking speech.

Given all of this, we should be very concerned about the apparent fact that our moral culture appears to have grown increasingly enamoured with casual debunking, or debunking that rests merely on some internal sense of obviousness. The more I think about this, the more I am struck by the extraordinary popularity that debunking enjoys, and one sign of its increasing popularity is that language itself has changed to suit the needs of debunkers. Consider: how often, in moral or political discourse do you hear someone claim that someone's action or utterance is "really about" X? How did this simple term "about" so quickly come to mean "the hidden motives or forces behind what someone says or does"? Moreover, we are becoming so good at debunking that entirely new terms are regularly minted which function as single-word debunkings. An example which clearly brings out the relevant epistemic hazards is the debunking term "snowflake."

This, as many readers may know, is a recently popularized term which, all on its own, implies that some target or another is operating under a hidden and problematic set of psychological causes. It has even entered the *Oxford English Dictionary*:

Snowflake(n): An insult to describe someone who is "overly sensitive or as feeling entitled to special treatment or consideration."

“Snowflake,” like all debunking terms, is a thick ethical concept: it is not just that people are sensitive, it is that they are *overly* sensitive (Kirchin 2013; Williams 1985). Moreover, when it is deployed in response to some action or utterance, it plainly aims to *debunk* that action or utterance by implying that it is the product of oversensitivity.

Now, debunking terms obviously *can* successfully refer. Surely there are snowflakes, just as there are mansplainers, bleeding hearts, and so forth. But once we work through the actual social uses of such terms, we will find that users often show little to no concern for basic epistemic standards. “Snowflake” is a paradigm case here, as an example will show.

Suppose that some students are calling for a university professor to be fired, and, as seems to always be the case, there are readily available videos showing the students angrily calling for her firing. The reasonable things to conclude, on the basis of this evidence, are:

- a. The students believe that the firing is morally required
- b. The students are angry at the professor, and possibly at the institution that employs her
- c. The anger is a primary cause of their expressed belief in (a), that the professor should be fired.

At this point, no normative judgment is implied by the evidence, and uses of the term “snowflake” are entirely inappropriate. As I insisted earlier when arguing against MacIntyre, there is no plausible principle which takes us from the fact that a moral belief is caused by anger to the conclusion that it is unjustified or suspicious. This is because anger can itself be caused by the prior perception of injustice; indeed, it is a perfectly *fitting* response to perceived injustice, given the harms or violations that injustice normally entails. So, what we need is:

- a. The anger is likely not caused by any perception of injustice, rather, it is heavily overdetermined: in virtue of their emotional sensitivity, these students will get angry at pretty much anything.¹¹ In other words, they are *oversensitive*, since there is not actually that much to get reasonably angry about.

This final complex claim could easily constitute a psychological debunking, and it is pretty much what the term “snowflake” means. It is clear that an emotional response might be seriously unreliable, or insensitive to the truth, if it will “go off” for pretty much any reason. Moreover, we know that there are people who cannot control their anger for various reasons, who are constantly *looking* for reasons to get angry, so the debunking hypothesis is not intrinsically unrealistic.

However, we should now back up and ask: what positive evidence do we have that (d) is true, that these students are habitually angry, rather than just angry in this one case, and how likely are we to render that evidence salient to our intended audience? This might seem *obvious*, but recall that a sense of obviousness cannot be sufficient, here. Do we have a long track record of their particular political activities and responses? Do we have any testimony of friends and colleagues concerning their character? Can we see the phrase “anger issues!” hastily scribbled on a therapist’s notepad? We have none of this. From a psychological standpoint, we are only justified in believing (a) through (c). The causal hypothesis contained in the use of “snowflake” is under-supported by the evidence we have.

Importantly, it is not at all impossible to acquire a decent amount of evidence for a psychological debunking. For example, if we want to debunk the actions or utterances of a putatively racist judge, we can simply examine their conviction record. Barring any confounding factors, a judge who sentences black defendants to 50% more jail time than other defendants for the same crimes is almost certainly harboring some unacknowledged racist attitudes and is ripe for a psychological debunking. Moreover, in such cases the moral importance of unmasking the judge is extreme, and easily outweighs the moral

hazards involved in attacking them personally. So I cannot be accused of setting an unreasonably high bar for a debunker to clear; it *can* be cleared, it is just that so often we simply do not even try to clear it.

Moreover, to return to the Simion-Lewis point about the *moral* content of an assertion, where is the contextually relevant *explanation* for the normative claim embedded in the use of “snowflake”? That is, where is the justification for the claim that students are overly sensitive? Even granting that they are much more sensitive than the average person, why is this *too* sensitive? Isn't it possible that they live in a world full of injustice and evil, and that it is the statistically average emotional sensibility that is seriously flawed?

Now, I can't resist pointing out that there is an easy way to motivate this kind of normative conclusion about the students. This is by thinking that *there is no injustice to perceive*, that the professor in question did nothing wrong. This belief almost trivially implies that the students are overreacting; if there is nothing to react to, then any reaction is an overreaction. But this is now a claim about the *truth* of the students' moral beliefs, and so we no longer have a psychological debunking on our hands. This is simply an expression of first-order moral disagreement, one which can only be settled by returning to the first-order moral reasons for and against the professor's behavior. So, the existence of causally efficacious anger (as implied by the term “snowflake”), or indeed of any other causal history whatever is now completely superfluous, a distraction at best. We must now do what Nagel and Cohen insisted we should have been doing all along, which is to stop trying to humiliate people with speculative psychology and to try to articulate our *reasons* for thinking the college professor did nothing wrong.

Thus far, the take-home message is this: psychological debunking is rarely accompanied by the kind of evidence that can ground justified belief in its causal hypotheses or its evaluative judgments. However, suppose one is in possession of a justified belief concerning a genuinely disreputable psychological influence, and that one can easily render the normative explanation salient to one's audience. Is one morally justified in enacting a debunking? Not necessarily; after all, given that you are choosing to undermine or humiliate someone, you have to be sure that you are doing so with a valuable purpose in mind.

3.2 | Evaluative hazards

What is the *point* of psychological debunking? What valuable state of affairs can help to justify our undermining the beliefs, utterances, or intentions of others by claiming that those states possess some disreputable causal history? Right off the bat, let us admit that debunking is often deployed as a point-scoring rhetorical tactic, and nothing more. We often aim to simply appear dialectically superior and to infuriate an opponent who is frustratingly unwilling to come around to our point of view. These aims, I take it, are not obviously worthwhile on their own, and few would argue that we are ultimately justified in humiliating another person for these reasons alone. In a way, this paper could be much shorter than it is: I could simply point to this depressing fact and suggest that we can be better.

But this does not mean that psychological debunking is intrinsically problematic. Indeed, so far as I can tell, there are at least very two worthy goals that can legitimately motivate a psychological debunker.

Individual Reform: no agent wants to be operating under the influence of such causal forces, and it is better, by the lights of any existing agent, to *know* that they are. Otherwise, they simply won't be able to reform themselves, and it is this reform which can justify a debunking.

Social Knowledge: when there are people in a society whose beliefs, values, and actions are distorted by negative causal forces, it is better, *ceteris paribus*, that other people be made aware of this fact.

These goals line up, rather neatly, with the two main contexts in which a debunking speech-act normally takes place. We sometimes directly inform individuals that they are “under the influence,” so to speak, and it is clear that this can be justified if this allows them to initiate valuable individual reform. In addition, we also undermine certain individuals *to* a group of other individuals, and this is potentially a means to valuable social knowledge, regardless of whether the target is present or listening.

Here, we can really appreciate the seriousness of the interpretive hazard. *If* we do not actually know that a person is “under the influence,” then we have no valuable knowledge to spread to our society, and we will be committing testimonial injustice by wrongfully reducing the social credibility of our targets (Fricker 2007). Moreover, we will be trying to provoke individual reform in a person who may not need it.

If, for example, student activists are angry because they genuinely perceive an injustice, we will be (a) spreading false information about their psychology, thereby (b) committing testimonial injustice by causing them to be seen as less credible when they are not, (c) potentially provoking them to “re-form” themselves, when (d) it is vitally important that they do *not* reform their emotional sensibility, which is, in this case, perfectly tuned to the existence of injustice (Lorde 1984).

This aside, under the assumption that we *do* know about some disreputable causal-psychological history, we might reasonably aim at Social Knowledge and Individual Reform, and our debunking could easily be justified by these ends. We may be doing some morally necessary dirty work, humiliating an individual for their own good or for the good of society in general. However, there is still the possibility that our means-ends beliefs are mistaken or unjustified. So, we need to ask: how often are we in a position to know that *debunking*, and not some other strategy, is an effective means to these potentially quite valuable ends?

3.3 | Means-ends hazards

It is not hard to see how this is so in the case of Individual Reform: persons are famously resistant to believing that they are under the influence of some hidden distorting factor. And while persons can surely be *too* sensitive and while we shouldn't preclude criticism of (for example) their ‘fragility’ in the face of criticism, it is nonetheless difficult to see how a population of human beings could be *completely* open-minded about such matters. This, again, is because psychological debunking is normally a form of humiliation.

To be humiliated is to lose social standing with respect to some standard that you care about. Ordinarily such a failure must reflect badly on your character or identity. The ordinary response to humiliation is shame, since shame, as an emotion, plays a vital role in maintaining our drive to meet those socially salient standards. But since shame is painful, there will *always* be some motive to avoid it in any person whose self-esteem is at all tied to their social standing (i.e. almost everyone). And one simple, well-worn avoidance strategy is to deny that one has deservedly lost social standing; in other words, to become defensive and to refuse to reform. In the case of a debunking speech-act, its target can try to avoid shame and humiliation by simply refusing to accept either the psychological or the moral claim implicit in the debunking. This implies that there will virtually *always* be some kind of resistance to debunking even when it meets the epistemic standards above.

This in turn means that simply *launching* a debunking story at someone in the public arena is often going to be a highly inefficient way to promote Individual Reform. Rather, more therapeutic methods will often be far more effective. For example, rather than outright express the *belief* that a person is “under the influence,” we might merely raise the possibility. Moreover, when possible, we might first

seek to establish the bonds of trust and friendship which can render such suggestions much more effective.

Of course, Individual Reform is not the only valuable goal we may have in mind when offering a psychological debunking. As I have mentioned, we may also aim to spread valuable social knowledge to others. But in the case of Social Knowledge, we may often overlook the fact that the activity of psychological debunking is, in a certain sense, deeply destructive to the preconditions for social knowledge as such. I will return to this point just below, but for now, let me just flag that we are not always right to think that we can best spread this form of knowledge via the pointed humiliation of a particular person or set of persons.

Let's pause to take stock. We have encountered a cascading series of conditions which must be met for a debunking speech-act to count as morally justified. First, we must avoid interpretive errors by ensuring that our assertion meets certain epistemic standards. Second, we must ensure that we are aiming to produce valuable states of affairs such as Individual Reform and Social Knowledge, and third, we must ensure that the debunking speech-act really is an effective means to those valuable ends. We must now turn to our fourth and final moral hazard: the possibility of unintended bad consequences.

3.4 | Unintended bad consequences

I lack the space to canvas all possible types of negative effects which might result from the practice of psychological debunking, so I'll just focus on two salient categories. The first concerns the deleterious effects on the debunker; the second concerns the harm done to our shared social life.

3.4.1 | The character of the debunker

Someone who makes a habit of attacking other persons may tend to damage their own character and live a socially impoverished life. A bully is often both a malignant and a tragic figure for this reason; not only do they harm others, they habitually encourage in themselves a shallow and alienated mode of being with others. It seems plausible to suggest that someone who makes a habit out of psychological debunking can produce the same result.

After all, and this is the point that motivated Nagel and Cohen, we very often have a choice. We don't normally *have* to impute disreputable motives or dark psychological forces to those we interact with, rather, we can engage substantively with the first-order reasons they have for their belief or attitude. To return to the case given at the outset of this paper, it is clear enough that the concerned third-party can simply say: "I don't think mildness and humility actually are good strategies in this company; let me tell you why." This strategy, while also undermining in some sense, has the advantage of highlighting mutually accessible evidence, reasons that can be in plain sight to all the conversational interlocutors. By contrast, the imputing of hidden dark motives does not often have this feature, and this is precisely what encourages the epistemic problems highlighted above.

In addition, we should all be wary of the fact that debunking is undeniably *fun*, that there is (for most people) a certain joy in knocking an interlocutor down a few pegs in this manner. Nietzsche plainly experienced a certain transcendent ecstasy each time he wielded his hammer: "what a pleasure for me!" he declares in this paper's epigraph. Yet, I don't think it should be controversial to say that when humiliating others becomes *enjoyable*, any sane moral view is going to tell us to worry about that.¹² Finally, we should be wary of this felt pleasure encouraging us to post-hoc rationalize our debunking activity, to ignore its real costs. Just as it shouldn't surprise us to learn that people who enjoy

alcohol are likely to under-estimate its objective risks (Greenfield and Rogers 1999), the fact that debunking is often fun can lead us to ignore its moral hazards. Though I am in no position to say that any reader of this paper is under this particular influence, this is surely a troubling possibility.

3.4.2 | Debunking and social warfare

We often think of tools in the following way: human beings have some purpose, and so they invent some tool to help them achieve that purpose. However, this is not always the full story: sometimes, human beings develop some tool which then suggests some further purpose, and which thereby produces some desire which would not otherwise exist.

For example, an overly simplistic history of the gun would go like this: people wanted to be able to exercise lethal force over large distances, so they invented guns. The sheer efficacy of the tool, we might think, is what caused it to become widespread. While this is surely part of the picture, the full story would have to include the fact that the existence of handguns encouraged more and more people to *want* to exercise lethal force over large distances, and this, in turn, encouraged the development and proliferation of more (and more lethal) guns. The very idea of such force would rarely even occur to a great many people unless instruments existed which could accomplish the purpose. Thus, a mutually reinforcing set of causes is perpetually in motion, where the tool and the corresponding desire symbiotically reinforce one another (Hodder 2014). This is the wisdom contained in the oft-repeated dictum from Maslow quoted at the outset of this paper. While sometimes we see a nail and pick up a hammer, other times we have simply grown so attached to our hammer that we see nails everywhere.

Many people are beginning to worry about the destructive effects of social fragmentation. While this sometimes gets called “polarization,” this is not the best term, for it suggests a one-dimensional fracturing of the civic body. But social fragmentation is in fact a much more complex phenomena, occurring along many dimensions, and it is driven by a populace that increasingly seeks to demonize or dehumanize those who do not share their moral or political standpoint on some particular issue. While people in a fragmented society will probably share similar positions relative to the rest of humanity, it remains the case that the activities of the other “side” provoke powerful emotional responses, driving them to demonize and dehumanize. This fragmentation is fairly well-documented in certain countries, though its precise contours are hotly debated.¹³ What is generally known is that anger and resentment—when bound up with in-group thinking and deep personal identification with political positions—are key causes. Here is Liliana Mason, summarizing a great deal of research:

Research in social psychology has found that when one of our social groups is in conflict with another group, that group identity becomes more central to our idea of who we are. The more intense the competition between the parties, the more we respond to politics as members of partisan teams, not as citizens. A threat to our groups’ status also causes us to think, feel, and act defensively... When the prestige of one of our groups is under threat (something which happens constantly in partisan politics), we—often unconsciously—lash out at our opponents... To a very real extent we are also being driven by an automatic, basic need to defend our social group. The result is more prejudice, activism and anger.

(Mason 2015, 58)

Of course, perhaps in some contexts what is needed is more anger, more activism, perhaps even more prejudice. For example, in the face of police violence and brutality, it is of course more than appropriate

for an African-American community to erupt into a kind of partisan activism and to lash out at those responsible for their suffering. My only point is more general: to the extent that social fragmentation is problematic in some context or another, we have yet another reason to worry about psychological debunking in that context. This is because someone looking to perpetuate a culture war could hardly *design* a better tool than the explanation of opposing views in terms of illicit causes rather than in terms of reasons.

First, psychological debunking in the political realm usually involves the imputation of bad motives or psychological traits to a target group, and it thus begins by separating the social world into an in-group and an out-group. Moreover, motives like envy, bigotry, selfishness, hypocrisy, and closed-mindedness are commonly ascribed by such acts of debunking. In a meta-analysis of 50 years' worth of data, Iyengar et al. discovered that the average American's willingness to impute these very motives to persons affiliated with a different political party rose by an extraordinary amount—a factor of between five and twenty, depending on how it is measured—over that timespan (Iyengar et al. 2012). This is what snowballing social warfare looks like, and it cannot be much of a stretch to suggest that habitual psychological debunking has been a major part of that story.¹⁴

What we have, then, is a tool that is uncannily ideal for a purpose: the perpetuation of social fragmentation. And since its proliferation can probably be explained by its reliably achieving that purpose, we may even be justified in concluding that one of its primary social functions is to produce social warfare.¹⁵ According to the overly simplistic view of tool-use outlined earlier, this shouldn't concern anyone, rather, those who are worried about social fragmentation should focus on the problematic desire to perpetuate social warfare. But, again, this is not how we should think about tools. It is more plausible to say that the desire to perpetuate social warfare would not be spreading so rapidly if this kind of social tool were not so readily available and so very functional. Even if it is a tool which efficiently accomplishes worthy goals—and I have already argued that this is only sometimes the case—those worthy goals must be balanced against the apparently extreme levels of fragmentation that psychological debunking is custom-tailored to create and reinforce.

A final word on social fragmentation. Recall that debunking, as an activity, can often only be justified by the valuable *social* knowledge it produces. Assuming that a person or group of persons is operating under problematic psychological influences, and assuming that we know this, we aim to *tell* our society about this fact in order to render our target less dangerous. However, one of the most troubling consequences of social fragmentation is the undermining of shared *epistemic trust* (Brady and Crockett 2019; McCraw 2015). This happens when the ordinary channels through which social knowledge is spread become blocked by affective resistance to knowledge-claims on the part of a large portion of the population.

Now, sometimes this can be valuable, inasmuch as harmful or oppressive relations of trust may be broken down.¹⁶ But if our goal is the spread of social knowledge, we must confront the fact that fragmentation destroys epistemic trust, and as such, even true debunking stories may not gain the kind of social traction they need to become part of “received wisdom” and to provoke real social change. At best, debunkers will only convince those who antecedently believe something very much like the debunking story already, and no social knowledge will be spread. This is one final way in which psychological debunking—in virtue of its tendency to increase affective fragmentation—can ironically destroy the preconditions for its own moral justification.

4 | CONCLUSION

Let me summarize. There are several reasons to worry about psychological debunking, which can easily counterbalance any positive reasons that may exist in its favor:

1. It is normally a form of humiliation, and we have a presumptive duty to avoid humiliating others.
2. It is all too easy to offer such stories without acquiring sufficient evidence for their truth,
3. We may aim at no worthy social or individual goals,
4. The speech-act itself may be a highly inefficient means for achieving worthy goals, and
5. We may unwittingly produce bad consequences which strongly outweigh any good we do achieve, or which actually undermine our good aims entirely.

As you may have guessed, I do not think that these problems exist in isolation; I believe that they are mutually reinforcing. For example, debunking stories would not augment social tensions so rapidly if debunkers were more likely to provide real evidence for their causal hypotheses. Moreover, if we weren't so caught up in social warfare, we'd be much less likely to ignore the need for evidence, or to ignore the need to make sure that the values which drive us are both worthy and achievable. It is my sincere hope that increased awareness of this vicious cycle can help us to break it. Nietzsche's hammer might be a joy to use, but it's time to learn to put it down.

CONFLICT OF INTEREST

None to report.

ACKNOWLEDGEMENTS

For helpful discussion, the author wishes to thank Kirun Sankaran, Regina Rini and Lauren Kopajtic. No thanks to his 1 year-old son, Luke William Smyth, without whom this paper would have been much more thoroughly thought-out and researched.

ORCID

Nicholas Smyth  <https://orcid.org/0000-0001-6155-1461>

ENDNOTES

- ¹ I use the technical term “expresses the proposition that ___” because not all debunkings are performed in so literal a sense. In the right context, a debunking can take a very simple form. If we have been discussing your lack of self-confidence, and you tell me that some recent writing of yours is terrible, I can express the proposition that your utterance is caused by your lack of confidence with a simple raised eyebrow and a sarcastic, “oh, yeah”?
- ² Of course, this three-part distinction may not be nearly as clean as it first appears. Socio-historical forces characteristically work through the psychologies of individual agents, and evolutionary histories very often make reference to social phenomena (as when, for example, evolutionary theorists explain various phenomena by citing *coalition games* or *reputations*). I set this worry aside as I do not think it bears on the project I undertake in this paper.
- ³ For example, most evolutionary histories of morality refer to the notion of *reputation within a group* in order to explain how altruistic behavior can evolve and stabilize in a population (Trivers 1971). It is very common for evolutionary theorists to claim that altruistic agents will flourish because other agents in the group hold their co-operative, reciprocal behaviour in comparatively high esteem. But it should be clear that this explanation is irreducibly social in character; the notion of a *reputation* is not one that simply falls out of, say, population-level genetics.
- ⁴ This “archeological” mode of interpretation has been wildly popular in the humanities for decades, and this surge in popularity has been ably criticized by Susan Sontag, Eve Sedgwick, Bruno Latour and Rita Felski (Felski 2015; Latour 2004; Sedgwick 1997; Sontag 1964/2001).
- ⁵ For example, Joshua Greene has alleged that deontological moral beliefs are influenced by untrustworthy emotional responses, and Shaun Nichols has argued that moral objectivists are under similar influences (Greene 2015; Nichols 2014).
- ⁶ Cohen and Nagel (1934), iv.

- ⁷ In giving his famous debunking argument against religious belief, Freud was careful to insist that a belief can be both illusory and true. And rather than argue that his psycho-social history of religion refuted the existence of God, Nietzsche explicitly says that “with the insight into that origin the belief *falls away*...” (Freud 1927/2012, Nietzsche 1886/1984)
- ⁸ Rini is unfortunately given over to what I call *merelyology*: rhetorically referring to empirical causes as “mere” empirical causes in a manner which implies (without explicitly saying) that they are somehow normatively suspect as such. Merelyologists write as though it is part of our shared argumentative background that causes cannot be reasons. For instances in this one paper alone, see (Rini 2020, 370, 371, 375, 376, 380, 381, 384). I should stress that I do not mean to pick on this author, this habit is endemic in the Kantian and moral-rationalist traditions, where references to such things as “mere desire” and “mere self-interest” abound.
- ⁹ It's worth noting that the desire to be both praised and praiseworthy is at the core of Adam Smith's moral theory in the *Theory of Moral Sentiments*, which is now enjoying something of a revival. See (Smith 1759).
- ¹⁰ These individuals are not just private individuals operating online, nor are they confined to the political right; some of them are strongly left-wing and many operate in academia. Foucault scholar Giorgio Agamben has suggested that the COVID-19 virus merely offered an “ideal pretext” for “scaling up exceptional measures beyond limitation” and a soothing balm for a populace that is increasingly seized by a “Look” (Agamben 2020, <https://www.journalpsychoanalysis.eu/coronavirus-and-philosophers/>). There is of course nothing wrong with highlighting the dangers of increased governmental control during a crisis, but to say that these drives to control and be controlled are the *only* operative motives here is to springboard into pure fantasy.
- ¹¹ This argument is just an application of a more general principle effectively defended by Katia Vavova. We cannot, Vavova argues, rest debunking on the mere possibility that someone's belief is off-track. We need, she says, *good independent reason* to think that it is so, and in this case what is needed is therefore some positive argument to the effect that the students are generally insensitive to moral truths (Vavova 2016).
- ¹² Of course, Nietzsche's own views don't count as “sane” on this view, but those of us who still cling to the allegedly “slavish” values of equality and peace must reject his view as dangerous and wrong-headed.
- ¹³ For an excellent and reasonably accessible introduction to the main debates, see *Political Polarization in America* (Sides and Hopkins 2015).
- ¹⁴ It is interesting to compare the deployment of more ordinary political slang with the activity of debunking. To deploy the term “pinko” is indeed to attempt to wound someone and to strongly suggest that they hold a morally reprehensible position. But it is not yet to say or imply that their belief-forming mechanisms *themselves* are unreliable, untrustworthy, or otherwise problematic. I may disagree with your position, but I have not yet implied that you *yourself* are morally or politically untrustworthy, that you were bound to arrive at a faulty or deficient position simply in virtue of being the kind of person that you are. This, I submit, is yet another way in which debunking is custom-tailored to erect powerful social barriers. Rather than simply suggest that a person's *position* is reprehensible, it can directly imply that a person isn't even worth talking to, that their cognitive and emotional faculties as such are not properly aligned. It is not hard to see that, in this respect, there is far more power in the term “snowflake” than in the term “pinko”: while the latter suggests that a person has arrived at a faulty position, the former implies that pretty much *any* position they arrive at is likely to be flawed, because that is just the kind of person they are.
- ¹⁵ That is, on the aetiological view of functions defended by many thinkers. See for example (Wright 1973; Cohen 1980; Millikan 1984).
- ¹⁶ For example, it is clear that African-American communities *ought not* to have trusted the police in Ferguson, MI. Their trust was a key part of the system that worked against them. Yet, the point remains: if we are to alert society at large to the existence of horribly biased individuals working in a deeply flawed system, we need them to trust us. Yet, in a society where debunking (rather than first-order argument) is the norm, that trust may be hard to secure.

REFERENCES

- Baier, A. 1985. *Postures of the Mind: Essays on Mind and Morals*. Minneapolis, MN: University of Minnesota Press.
- Boudry, M, F Paglieri, and M Pigliucci. 2015. “The Fake, the Flimsy, and the Fallacious: Demarcating Arguments in Real Life.” *Argumentation* 29 (4): 431–456.

- Brady, W. J., and M. J. Crockett. 2019. "How Effective Is Online Outrage?" *Trends in Cognitive Sciences* 23 (2): 79–80. <https://doi.org/10.1016/j.tics.2018.11.004>
- Cohen, G. A. 1980. "Functional Explanation: Reply to Elster." *Political Studies*, 3, 129–35. Accessed December 04, 2014. <http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=14399399&site=ehost-live>.
- Cohen, M. F., and E. Nagel. 1934. *An Introduction to Logic and Scientific Method*. Read Books Ltd.
- D'Arms, J., and D. Jacobson. 2000. "Sentiment and Value." *Ethics* 110 (4): 722–48.
- Felski, R. 2015. *The Limits of Critique*. Chicago, IL: University of Chicago Press.
- Freud, S. & 1920. *A General Introduction to Psychoanalysis*. New York, NY, US: Horace Liveright Publishing, 90–100.
- Freud, S. 1927/ 2012. *The Future of an Illusion*. New York: Broadview Press.
- Fricker, M. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*, Vol. 69. Oxford: Oxford University Press.
- Gilbert, P. 1997. "The Evolution of Social Attractiveness and Its Role in Shame, Humiliation, Guilt and Therapy." *British Journal of Medical Psychology* 70 (2): 113–47.
- Greene, J. D. 2015. "Beyond Point-and-Shoot Morality: Why Cognitive Science Matters for Ethics." *Law and Ethics of Human Rights* 9 (2): 141–72.
- Greenfield, T. K., and J. D. Rogers. 1999. "Alcoholic Beverage Choice, Risk Perception and Self-reported Drunk Driving: Effects of Measurement on Risk Analysis." *Addiction* 94 (11): 1735–43. <https://doi.org/10.1046/j.1360-0443.1999.9411173510.x>
- Hodder, I. 2014. "The Entanglements of Humans and Things: A Long-term View." *New Literary History* 45 (1): 19–36.
- Iyengar, S., G. Sood, and Y. Lelkes. 2012. "Affect, Not Ideology: A Social Identity Perspective on Polarization." *Public Opinion Quarterly* 76 (3): 405–31.
- Joyce, R. 2007. *The evolution of morality*. Cambridge: MIT press.
- Kahane, G. 2011. "Evolutionary Debunking Arguments." *Noûs* 45 (1): 103–25.
- Kirchin, S. 2013. *Thick Concepts*. Oxford: OUP Oxford.
- Korman, D. Z. 2019. "Debunking Arguments in Metaethics and Metaphysics." In *Metaphysics and Cognitive Science*, edited by A. Goldman, and B. McLaughlin, 337–363. Oxford: Oxford University Press.
- Latour, B. 2004. "Why Has Critique Run Out of Steam? From Matters of Fact to Matters of Concern." *Critical Inquiry* 30 (2): 225–48.
- Lewis, M. 2019. "The Norm of Moral Assertion: a Reply to Simion." *Ethical Theory and Moral Practice* 22 (4): 1043–9.
- Loorde, A. . "The Uses of Anger: Women Responding to Racism." *Sister Outsider* 127: 131.
- MacIntyre, A. 1981/ 2013. *After Virtue*. A&C Black.
- Mason, K. 2010. "Debunking Arguments and the Genealogy of Religion and Morality." *Philosophy Compass* 5 (9): 770–8.
- Mason, L. 2015. "Party Polarization is Making Us More Prejudiced." *Political Polarization in American Politics*, 55–60. New York, NY: Bloomsbury Press.
- McCraw, B. W. 2015. "The Nature of Epistemic Trust." *Social Epistemology* 29 (4): 413–30.
- Miller, W. I. 1995. *Humiliation: And Other Essays on Honor, Social Discomfort, and Violence*. Ithaca, NY: Cornell University Press.
- Millikan, R. G. 1984. *Language, Thought and Other Biological Categories*, Vol. 14. Boston, MA: MIT Press.
- Nichols, S. 2014. "Process Debunking and Ethics." *Ethics* 124 (4): 727–49.
- Nietzsche, F. 1887. *On the Genealogy of Morals: A Polemic*. Oxford: Oxford University Press.
- Nietzsche, F. W. 1886/ 1984. *Human, All Too Human: A Book for Free Spirits*. Lincoln, NE: University of Nebraska Press.
- Nussbaum, M. C. 2003. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press.
- Rachels, J. 1997. "The Moral Argument for Vegetarianism." *Can Ethics Provide Answers?*. 99–107. New York, NY: Rowan and Littlefield.
- Ricoeur, P. 2008. *Freud and Philosophy: An Essay on Interpretation*. New Delhi: Motilal Banarsidass Publisher.
- Rini, R. 2020. "Contingency Inattention: Against Causal Debunking in Ethics." *Philosophical Studies* 177 (2): 369–89. <https://doi.org/10.1007/s11098-019-01397-8>
- Ronson, J. 2016. *So You've Been Publicly Shamed*. New York, NY: Riverhead Books.
- Sauer, H. 2018. *Debunking Arguments in Ethics*. New York, NY: Cambridge University Press.
- Sedgwick, E. K. 1997. "Paranoid reading and reparative reading, or, You're so paranoid, you probably think this introduction is about you." In *Touching Feeling: Affect, Pedagogy, Performativity*, 138–149. Durham, NC: Duke University Press.

- Sides, J., and D. J. Hopkins. 2015. *Political Polarization in American Politics*. New York, NY: Bloomsbury Publishing USA.
- Simion, M. 2018. "The Explanation Proffering Norm of Moral Assertion." *Ethical Theory and Moral Practice* 21 (3): 477–88.
- Smith, A. 1759. *The Theory of Moral Sentiments*. London, UK: Dover Publications.
- Sontag, S. 1964/ 2001. *Against Interpretation: And Other Essays*, Vol. 52. London: Macmillan.
- Statman, D. 2000. "Humiliation, Dignity and Self-Respect." *Philosophical Psychology* 13 (4): 523–40.
- Tappolet, C., and M. Rossi. 2015. "What is Value? Where Does it Come From? A Philosophical Perspective." In *The Value Handbook: The Affective Sciences of Values and Valuation*, edited by T. Brosch, and D. Sander, 3–22. Oxford: Oxford University Press.
- Trivers, R. L. 1971. "The Evolution of Reciprocal Altruism." *The Quarterly Review of Biology* 46 (1): 35–57. <https://doi.org/10.1086/406755>
- Vavova, K. 2015. "Evolutionary Debunking of Moral Realism." *Philosophy Compass* 10 (2): 104–16.
- Vavova, K. 2016. "Irrelevant Influences." *Philosophy and Phenomenological Research* 96 (1): 134–52. <https://doi.org/10.1111/phpr.12297>
- Williams, B. A. O. 1985. *Ethics and the Limits of Philosophy*, Vol. 83. Cambridge: Harvard University Press.
- Wright, L. 1973. "Functions." *Philosophical Review* 82 (2): 139–68.

How to cite this article: Smyth N. A moral critique of psychological debunking. *J Soc Philos.* 2021;00:1–18. <https://doi.org/10.1111/josp.12402>