## Introduction

Kant identified three questions that philosophy must address: What can I know? What must I do? What may I hope? Today, a fourth question preoccupies many philosophers: Why was my paper rejected? Philosophers, and academics in many other disciplines, pore over reports from reviewers and editors, shaking their heads in disgust. Why didn't they recognize the brilliance of my ideas? Who thought *that* clown was competent to review my paper? I'm going to argue that answering these sorts of questions – questions concerning how reviewers evaluate the papers they consider, and what influences their decisions – requires us to get to grips with the factors that modulate intellectual charity.

The narrower goal of this Element is to shed some light on peer review, and in particular on some of the evidence for the widespread feeling that peer review is broken.[1] It will focus, in particular, on the evidence arising from the sometimes widely disparate responses competent readers might have to the same paper. A number of theorists have pointed to the fact that papers subsequently regarded as groundbreaking were rejected multiple times before finally finding a home, and to empirical evidence that accepted papers do not fare particularly well when resubmitted with cosmetic changes, and concluded that peer review doesn't do a good job at tracking quality. I will argue that though this evidence does show that what gets published is somewhat arbitrary, it does not show that journals don't do a reasonable job at selecting excellent papers.

The broader goal of the Element is to understand the difference that attitudes – those of readers, of reviewers, and of authors themselves – do and should make to the assessment of texts (whether journal articles or books). I will argue for a counterintuitive thesis: the quality of a paper (or a book: texts more generally) is not wholly intrinsic to it, but is a partly relational property – it depends in part on contextual factors. Context modulates the attitude we take to a paper, and that attitude, in turn, helps to shape our perception of its value. We can't use the history of a paper as a metric by which to assess its unchanging value, because its value is *not* unchanging before and after publication. Publication is one among many factors that (rationally) affect our perception of a paper's quality. Who wrote it, where it is published, what others have, or have not, said about it – all these, too, make a difference to our rational perception of its quality.

In Sections 1 and 2 of this Element, I will examine how intra- and extratextual features of texts affect reviewer attitudes toward them. In Sections 3 and 4, I will broaden the focus, to include the role that reader and author attitudes can and should play in modulating the perceived quality of a text. In Section 3, I ask

[1] The phrase "peer review is broken" returned about 321,000 hits on Google (as of December 11, 2022).

whether reviewers should be less trusting of the intentions of authors, in the light of recent attempts to hoax journals. I'll argue that such wariness is corrosive. Academic research is heavily reliant on trust, and we risk important goods if we become less trusting. In the final section, I will turn from the attitudes of reviewers and subsequent readers to the attitudes of authors, asking what attitudes we permissibly take to the claims we make in our own papers. I'll suggest that *assessor* attitude is partially a response to perceived *author* attitude: we permissibly modulate trust and intellectual charity by indicating our own degree of commitment to our claims (we may also mislead to the same effect, of course). In short, paper quality does not depend exclusively on the arguments presented but also on the attitude of assessors, and these attitudes can be modulated by many factors including the attitude the author is perceived to take to them.[2]

These claims – that quality is partially a relational property, that assessor attitude makes a difference to perceived quality, and that attitude is rationally modulated by a range of intra- and extratextual factors – are true well beyond philosophy. They're probably somewhat less true with regard to formal and deductive arguments than those which are inductive or abductive, but they play a role everywhere. This is a work *of* philosophy, and philosophy is the discipline I know best, so I will focus on it, but the issues I'll consider arise in other disciplines too. Wherever people are in the business of assessing one another's work, explicitly or implicitly (for example, by choosing to cite *this* paper, rather than *that*), they arise. I'll have little to say about these other disciplines, though: the implications will usually be clear.

It's important to emphasize that this Element is concerned with only a narrow slice of the evidence against peer review: the evidence that arises from the (erroneous) assumption that paper quality is a wholly intrinsic property. There are many concerns it leaves aside, and some may motivate the replacement of peer review by some other system. It's equally important, though, to recognize that assessor attitude will continue to play a role in perceptions of quality under any proposed replacement for peer review. So long as we're in the business of making distinctions on the basis of quality, the considerations I'll discuss will be relevant. Of course, assessing work on the basis of quality is something we can't avoid doing. We often need to settle how much weight to give to the claims of

[2] It bears noting that the basic view defended in this Element is in some ways parallel to, and may in fact be inspired by, the account of interpretation famously developed by Hans-Georg Gadamer (1975). To my knowledge, Gadamer never applied his framework to the kinds of issues I consider here. I'm not sufficiently versed in his views to assess the degree to which my account departs from his.

a book or paper, and answering that question inevitably involves assessing its quality.

Before beginning the exploration of the difference that attitude rationally makes to perceptions of quality, a brief explanation of the focus on peer review, and on its nature, might be in order. It's important, first, to emphasize the *centrality* of peer review, both to the professional lives of philosophers (and researchers in many other fields) and to the growth of knowledge. The stakes are high for individuals. The journal article is the principal unit of publication in professional philosophy (as in most, though not all, academic disciplines), and a solid record of publication in journals is almost always needed for professional success. Hiring decisions are made on the basis of candidates' publication records, and tenure requires continuing publication. Grants and self-esteem are also dependent on track record.

Peer review is also widely regarded as playing an important *epistemic* role. It is supposed to act as a guarantor that research was well-conducted, and that findings are reliable enough to be taken seriously. Anyone can publish their arguments and their research. Anyone can make a website, and if one wants one's writing to appear in journal format, there are plenty of predatory journals that will publish anything for a fee. But if one wants to be taken seriously, it's important to publish in a properly peer-reviewed journal.

The professional centrality and epistemic significance of peer review both reflect the fact that it's widely held to be a reliable measure of the quality of research. Peer review is usually *anonymous*. Review is usually either single or double anonymous. In single anonymous review, reviewers know the identity of the author(s) but not vice-versa; reviewers are therefore free to be honest without fearing retaliation or worrying about maintaining good relations with authors. In double anonymous review, reviewers and authors are both anonymous: this is often held to be superior to single anonymous review because it ensures that reviewers will not be swayed by the prestige (or lack thereof) of the author(s).

While peer review is widely held to be especially reliable at identifying quality, peer-reviewed journals are not all created equal. Some are more prestigious than others. Philosophers, like other academics, take the prestige of a journal into account when choosing where to submit their papers. Journal choice is guided by "fit" – different journals focus on different areas within philosophy, as well as taking different approaches – and prestige. Generally speaking, a philosopher will aim for the most prestigious journal they think might possibly take their paper. She will almost certainly be sensitive to journal prestige (within her subfield, or within the profession as a whole), regardless of whether she believes that prestige genuinely correlates with quality. The

unemployed and insecurely employed know that prestige publications boost their chances of secure employment and the untenured know that such publications may be required for tenure. Prestige publications are also required for securing grants. And, rightly or wrongly, almost all of us are motivated by a desire for recognition: the fact that many of their peers think more highly of work published in prestige journals motivates philosophers to try to publish in them.

But the more prestigious the journal, the lower the acceptance rate: the competition is intense and the stakes are high. While accurate and up-to-date information is hard to come by, the median acceptance rate in philosophy seems to be substantially lower than in most other fields (Weinberg, 2018). The acceptance rates for the most competitive philosophy journals, like *Ethics* (Driver and Rosati, 2021) and *Philosophical Review* (Philosophical Review, 2022) are lower than the rates for *Science* (Science, 2022) and *Nature* (Nature, 2021). Not only is the competition for the top spots more intense, the number of prestigious outlets is relatively small. Almost every respectable journal rejects most of the submissions it receives; many in philosophy reject upward of 90 percent of submissions. When the rejection is accompanied by one or more reviewer reports (desk rejection – rejection by the editors without peer review – is not unusual), we often read these reports obsessively. We may take to social media to slam them. There's a Facebook group called "Reviewer 2 Must be Stopped!," where academics go to vent about the perceived (and sometimes actual) ineptitude of reviewers (Reviewer 2 is the proverbially unfair, obtuse, rude, and often incompetent reviewer). The group has nearly 75,000 members, so we can be confident that obsession with bad reports is common across many disciplines. Sometimes reports are very helpful and illuminating (sometimes, they've even convinced me that the journal was right to reject my paper). Sometimes, they're incompetent. Mostly, they fall somewhere in between these extremes.

No one doubts that peer review is far from perfect. Papers that shouldn't have been published get through far more often than they should, and many, many good papers are rejected (whether that's a failing in need of rectification is one of the questions I'm concerned with here). There are now serious ongoing debates over its value. Peer review is now regarded as the gold standard for scientific credibility. Yet its dominance is surprisingly recent. *Science* (founded in 1880) introduced it around 1940; *The Lancet* (1823) only in 1976 (Shema, 2014). Prior to its introduction, journals relied on editors' opinions to select papers for publication. Science without peer review is certainly possible, and many researchers have proposed alternatives to the current system that would reduce or eliminate some of its failures (at minimum, by reducing or eliminating

the lengthy delays, and therefore opportunity costs, it involves). We might for instance adopt the model already dominant in some parts of physics: postpublication review (Heesen and Bright, 2020). This model abolishes journals altogether or places much less weight on publication in them. Instead, papers are uploaded to a preprint server, and it is their reception by the community of scholars that determines their success.

Perhaps we *should* replace peer review. I won't attempt to assess the issue here. In the end, I'm concerned not with peer review per se, but with epistemic issues in the assessment of the kind of work that gets peer reviewed, whatever the context in which that assessment occurs. I don't apologize for framing the issue around peer review, however, because (right now) it is in peer review that the most consequential assessment takes place, and because the considerations I'll highlight bear quite directly on some of the evidence for the claim that peer review is broken. Recognizing the difference that attitude can make to our justified assessment of a paper, and how attitude responds rationally to intra- and extratextual features beyond the argument and the data, can help to defuse some of the anger peer review arouses. In the concluding section, I'll argue that it also alerts us to some of the trade-offs, between different epistemic goods and between epistemic and nonepistemic goods, we face in designing any refinements to or replacement of peer review.

I'm going to begin in what might seem a surprising place: by plunging into bullshit, and the literature that discusses it. Section 1 will examine this literature, with an eye to understanding why we sometimes call bullshit on texts. Section 2 will apply the lessons learned to peer review. Understanding bullshit will provide us with tools that can be used to understand more nuanced and sympathetic assessments of philosophical work. By wading through the bullshit, we'll come to appreciate the decisive difference that intellectual charity routinely and inevitably plays in our assessments. These sections establish (to my satisfaction, at any rate) that attitude can make a decisive difference in our assessment of texts. Sections 3 and 4 inquire into the attitudes we justifiably take to texts, first as reviewers (Section 3), then as authors (Section 4).

## 1 Bullshit Philosophy

Why tread in bullshit? One reason might be because philosophy itself is often dismissed as bullshit. Perhaps understanding the nature of bullshit, or the disposition to attribute it, might help us to defend it against the accusation. I have a different goal in mind, however. I'm going to argue that understanding the conditions under which we're tempted to attribute bullshit sheds light on the ways in which our attitudes affect our assessment of one another's papers. The

attribution of bullshit to an assertion or a text generally arises from a lack of intellectual charity in reading. Reviewers rarely withhold intellectual charity to such an extent that they're tempted to see the papers they review as bullshit, but – I'm going to claim – much smaller differences in intellectual charity can and do explain much smaller (but still often decisive) differences in the assessment of papers. They may, for example, make the difference between "reject" and "revise and resubmit." We'll be better able to appreciate the smaller differences attitude makes, in peer review, for example, once we've seen how it can transform the tolerably clear into the incomprehensible.

The inquiry into bullshit will also guide us in identifying the kind of factors that modulate attitudes. Calling bullshit is, in paradigm cases, a response to both intra- and extratextual cues – in particular, cues that suggest the genre it belongs to. In subsequent sections, we'll identify other cues that play similar roles in modulating assessor attitudes.

## 1.1 Understanding Bullshit

The accusation that philosophy is bullshit is common (the phrase "philosophy is bullshit" returns more than 39,000 results on Google). Eminent scientists like Stephen Hawking, and influential science popularizers like Neil DeGrasse Tyson and Bill Nye have dismissed its value (Goldhill, 2016). Philosophers themselves are often ready to denigrate particular areas or approaches as bullshit, or something equivalent to bullshit. Dismissals of all or parts of "continental philosophy" as bullshit are not at all uncommon, though thankfully rarer now than in the past. The name-calling is by no means one-sided: Graham Harman is scarcely more charitable when he decries the "shallowness, false dichotomies, lack of imagination, robotic chains of reasoning, and the aggressive self-assurance that typifies analytic philosophers at their worst"; lest you think he reserves his scorn only for some pockets of analytic philosophy, he adds "we need to question the assumptions of this entire school" (Harman, 2009: 167–168). But the accusation that a school of philosophy is bullshit is reserved for "difficult" and "obscure" work. Even if we think that the kind of philosophy characterized by numbered premises, formal proofs, and attempts at exceptionless definitions is arid and sterile, we're unlikely to dismiss it as bullshit.

Is (some) philosophy bullshit? Harry Frankfurt (2009) has influentially argued that bullshit consists in assertions made without regard to whether they're true. If that's what's meant when some people dismiss some schools of philosophy as bullshit, it's often question-begging: it begs the question against those schools that aim at bringing us to see the world differently, rather

than at stating truths about it. It might be true that some philosophers don't care whether what they're saying is literally true, but it doesn't follow – without a lot of argument – that their work deserves to be dismissed as bullshit. Just as importantly, sincerity doesn't guarantee that an assertion isn't bullshit. Think of the accusation "he believes his own bullshit"; if Frankfurt-style bullshit was the only kind there was, that accusation would make little sense.

Frankfurt's account also doesn't capture how "bullshit" is used in the empirical literature. This literature builds on Gordon Pennycook and colleagues' (2015) path-breaking work. They define "bullshit" in terms of form and content, rather than the attitudes of those who produce it: on their account, bullshit consists in assertions that, in virtue of their syntactic structure, seem to convey meaning but (supposedly) lack genuine content. Their paradigms are the obscure sayings of the purveyors of supposed new age wisdom, like this example they draw from Deepak Chopra: "Attention and intention are the mechanics of manifestation" (Pennycook *et al*., 2015: 550). Their interest as psychologists is in what characteristics of hearers or readers make them receptive to these assertions; that is, what are the psychological correlates of finding bullshit profound?

Pennycook *et al*. take their account of bullshit to pick out assertions that also count as Frankfurt-bullshit, but they're wrong about that. Deepak Chopra may or may not be indifferent to whether his assertions are true. His attitude is irrelevant to whether they *are* bullshit, on their account. Gerry Cohen (2013) has offered an alternative account, which seems better able to capture what Pennycook *et al*. call "pseudo-profound bullshit." Cohen doesn't aim to replace Frankfurt's notion; rather, he argues that we need more than one account to capture all the bullshit that's out there. Cohen-style bullshit can be sincerely intended; its distinguishing mark is that it's "unclarifiable" (Cohen, 2013: 104).

Cohen also brings the discussion directly back to philosophy: he singles out "Francophone philosophical culture" as the "most successful producer of bullshit, both in respect of the volume of bullshit that it has produced and in respect of the warmth with which that bullshit has been received" (Cohen, 2013: 108). Again, Cohen doesn't dispute that those who produce the work he mentions (Deleuze, Derrida, Kristeva, and Lacan, as well as the Althusserian Marxists who are his special target) are sincere. He thinks they aim at truth, at least truth of a kind. Nevertheless, what they say is *not* true: it's not even false.

Cohen's original interest in bullshit stems from an avowedly autobiographical motivation. In his twenties, he tells us, he read a great deal of French Marxism, most of it stemming from the Althusserian school. He goes on to report that he struggled to extract meaning from these texts, and when he was able to make some sense of them attributed more significance to the claims than

they really deserved, in virtue of the effort he'd expended extracting them. His dalliance with Althusserian Marxism subsequently made him intolerant of bullshit, and when he founded a Marxist discussion group he called it the Non-Bullshit Marxism Group (Cohen, 2013: 95). We'll return to Cohen's confessed intolerance of bullshit; I think his preconceptions about these texts matter a great deal to his responses to them. In any case, his principal example – his "wonderful example of bullshit" – comes from a former student of Althusser's, Étienne Balibar: "This is precisely the first meaning we can give to the idea of dialectic: a logic or form of explanation specifically adapted to the determinant intervention of class struggle in the very fabric of history" (Balibar, 2014: 97). Cohen might once have assessed Balibar's assertion as profound; he now dismisses it as bullshit. Can the psychological work shed light on why he was once impressed, and perhaps on why he is now dismissive?

Pennycook *et al*. (2015) asked their participants to rate the assertions they were presented with – drawn from Deepak Chopra's tweets and from two online new age bullshit generators – for profundity, alongside motivational quotes and mundane assertions. They found that bullshit receptivity was correlated with an intuitive cognitive style and lower cognitive ability, as well as with higher rates of religious and paranormal belief. Allegedly, people are impressed by bullshit due to their failure or inability to detect the vacuity of the stimuli. I think it's vanishingly unlikely that Cohen fits their model, in cognitive style, in ability, or in beliefs in the supernatural. Perhaps another account of bullshit receptivity might do better at explaining his dalliance with Althusserian Marxism?

The empirical literature can be mined for other explanations. Pennycook *et al*. (2015) probed the interaction between the (allegedly nonsensical) content of the assertions that feature as stimuli in their work and the dispositions of those who are receptive to them. Other work looks instead to the apparent *source* of assertions (as well as their content) to explain receptiveness. Dan Sperber (2010) has identified the *guru effect*, whereby high credibility is assigned to an assertion in virtue of the fact that it is obscure *and* it is believed to stem from someone who is a "guru"; that is, a source of wisdom. Sperber takes the guru effect to partly explain the success of some philosophers, "especially but not uniquely in the so-called continental tradition" (Sperber, 2010: 587); Sperber goes on to quote Sartre, Heidegger, and Derrida; perhaps Balibar could be added to Sperber's list.

More recently, the guru effect has been joined by the *Einstein effect* (Hoogeveen *et al*., 2022). In this study, stimuli generated by the New-Age Bullshit Generator[3] were attributed to either a scientist or a spiritual leader.

[3] http://sebpearce.com/bullshit/.

Across more than 10,000 participants in 24 countries, assertions were judged as more credible when attributed to a scientist than to a spiritual leader; even religious participants tended to see statements attributed to a scientist as more credible than identical statements attributed to a spiritual authority. Perhaps the success of bullshit in Francophone philosophy is due to the special status of philosophers in France: because they have the status of "gurus," their nonsense is attributed with a great deal of credibility, despite – or perhaps in virtue of – its "unclarifiability."

The source and content effects just listed – the guru effect, the Einstein effect, the effects of bullshit on perceptions of profundity – are typically conceived as different ways in which assertions are given an underserved boost in credibility in the eyes of those who are susceptible. The credibility inflation is held to be undeserved, because – after all – the statements that are rated as credible and even profound are designed to be *meaningless*. The bullshit receptivity scale developed by Pennycook *et al*. (2015: supplementary material) uses items like the following: "We are in the midst of a high-frequency blossoming of interconnectedness that will give us access to the quantum soup itself." The Einstein effect was measured using stimuli like: "Yes, it is possible to exterminate the things that can confront us, but not without hope on our side. Turbulence is born in the gap where transformation has been excluded. It is in evolving that we are re-energized" (Hoogeveen *et al*., 2022: supplementary material). Since these are bullshit statements, the assignment of *any* degree of credibility and profundity to them is undeserved.

Of course, credibility boosts in virtue of these kinds of effects might often attach to genuinely meaningful statements too: they might be perceived as more profound than they deserve, given their relatively banal contents. Meaningful statements can have their credibility boosted in still other ways. Several studies have found that the use of irrelevant neuroscientific information boosts the credibility of assertions for naïve participants (Weisberg *et al*., 2008; Fernandez-Duque, Evans, and Hodges, 2015). There's no reason to think that these kinds of effects are limited to extraneous neuroscientific information. Scientific jargon of all sorts probably impresses many people.

Cohen may have briefly fallen victim to the guru effect or the Einstein effect: the status of the philosopher within France may have led him to take Balibar's claims to be more significant than they deserve. In doing so, he would have fallen victim to what I'm calling an *extratextual* influence: his assessment of the text was influenced by awareness of facts outside the text. Alternatively, he may have been moved by *intratextual* factors: for example, the obscurity of the

claims. Of course, these kinds of influences interact and are rarely cleanly separated. The content of a new age aphorism suggests a "spiritual leader" as its source; conversely, banal assertions from a scientist get little or no boost in credibility in virtue of their source. One way or another, Cohen was taken in by bullshit, or so he later came to think.

## 1.2 Deserving Bullshit

With these tools in hand, we can now put fine clothes on our resentments. We analytic philosophers, with our careful arguments and our rigor and our devotion to clarity, are outshone in the public eye by continental philosophers who disguise their essential triviality under layers of apparently significant pseudo-profundity. Their claims, when they bother to make them, are either devoid of meaning altogether, or cover essentially mundane thoughts with an unearned glamour (Shackel, 2005). We may not be able to partake of their aura, but at least we have the satisfaction of exposing them for what they are, and laying bare the tricks that take in the unwary.

None of this is true. Whatever the merits of continental philosophy – or particular strands within it – the empirical work on bullshit does not provide us with any grounds for dismissing or downgrading it. It might be true that the work of (some) continental philosophers receives a credibility boost in virtue of features of their content that might be paraphrased away without loss of meaning, or in virtue of the status of the intellectual, and specifically the philosopher, in France. But it's not true that any such credibility boost is strictly extraneous or necessarily undeserved. It may be fully rational to assign a higher degree of credibility to a text (with "text," understood broadly, to encompass anything from an assertion through to a corpus of work) in virtue of features of its content that could be paraphrased away, or of its apparent source.

This is clearest with regard to the Einstein effect. It should be entirely obvious that we do and *should* assign higher credibility to assertions by scientists speaking in their domain of expertise, in virtue of the fact that they issue from scientists. We're rarely in a position to verify scientific claims for ourselves; instead, we accept the claims of specialists speaking within the domain of their expertise because we take them to be much better positioned to make such claims than we are. It's worth emphasizing how routine this kind of deference is, especially in light of the politicization of science that was a central feature of the COVID-19 pandemic (and which led many people to doubt the epistemic credentials of epidemiologists). We may not automatically accept the diagnoses of medical professionals, but we almost always assign high credibility to them,