Intellectual Virtue Signaling and (Non)Expert Credibility

ABSTRACT: *In light of the complexity of some important matters, the best epistemic strategy for laypersons is often to rely heavily on the judgments of subject matter experts. However, given the contentiousness of some issues and the existence of fake experts, determining who to trust from the lay perspective is no simple matter. One proposed approach is for laypersons to attend to displays of intellectual virtue as indicators of expertise. I argue that this strategy is likely to fail, as non-experts often display apparent intellectual virtues while legitimate experts often display apparent intellectual vices. Then, I argue that this challenge is difficult to overcome, as experts who attempt to better exhibit apparent intellectual virtues would likely compromise their own reliability in the process. Finally, I discuss two conclusions—one more optimistic and one more pessimistic—that one might draw concerning the role of intellectual virtue in the identification of experts.*

KEYWORDS: autonomy, experts, intellectual virtue, open-mindedness, virtue signaling

Introduction

In many domains of inquiry, the most truth-conducive strategy for laypersons is to defer to relevant experts, at least when there is a consensus among them. This point underscores the importance of laypersons being able to identify the relevant experts in various domains. This is a notoriously difficult task, especially in those cases in which there exists substantial controversy about target questions. Moreover, allegations of conspiratorial activity on the part of putative experts may complicate laypersons' identification of reliable sources.

In section one, I recount strategies that have been identified to facilitate laypersons' identifications of relevant experts. In the process, I offer some critical remarks concerning some of the proposed strategies. Then, in section two, I draw attention to a neglected complication in the identification of experts. In short, non-expert sources often outperform legitimate experts in displays of intellectual virtue. Insofar as laypersons rely on displays of intellectual virtue as credibility cues, they can be expected to misallocate trust. In section three, I consider the extent to which the problem identified here can be solved. I argue that the challenge is severe, as attempts by experts to increase their perceived credibility by signaling intellectual virtue would either be insincere and often counterproductive or would reduce their effectiveness *qua* experts. This is because good intellectual

practice often conflicts with what, from the lay perspective, *seems to be* good intellectual practice. I conclude with some brief remarks on how this challenge might in principle be met.

1. The Identification of Experts

Following Alvin Goldman, I define an expert in a domain as:

> [S]omeone who possesses an extensive fund of knowledge (true belief) and a set of skills or methods for apt and successful deployment of this knowledge to new questions in the domain. (2001: 92)

So construed, expertise is an objective matter of actual and potential reliability. It is thus possible, on the present approach, that those widely regarded as experts are not legitimate experts. For example, it is in principle possible that an astrophysicist with an excellent reputation, credentials, and professional positions to his name is only an apparent expert. Likewise, it is possible in principle on the present approach that some individuals lacking the trappings of expertise are experts nonetheless. How regularly the trappings of expertise are misaligned with actual expertise will depend on the integrity of the processes by which reputation, credentials and the like are awarded, individuals' opportunities to pursue the trappings of expertise, individuals' interest in doing so, and so on. Thus, for some examples, nepotism and other forms of favoritism, as well as discriminatory practices, can cause misalignments between legitimate expertise and the markers of expertise.

Because expertise is an objective matter, and actual expertise may in principle depart from perceived expertise, the identification of legitimate experts is a non-trivial task. In some cases, the identification of legitimate experts is not necessary for truth-conducive epistemic reliance on others. For example, the proposition that Finland exists is, at least for most people, believed to be based on the testimony of others. However, because there is virtually no disagreement concerning this proposition,[1] truth-conducive social epistemic reliance on others does not require distinguishing between experts and non-experts. For a broad range of other questions, however, there exists substantial disagreement that complicates the epistemic policy of deference to experts. Notably, while many social epistemologists have focused on cases of disagreement among putative experts bearing the trappings of expertise—disputes between professional scientists, for example—the challenge is broader than this. In addition to disagreements within epistemic institutions, propositions that are widely accepted within such institutions are often challenged from without. For example, proponents of alternative medicine regularly challenge orthodox medical claims and practices. The challenge facing the layperson is thus not merely to select sources among putative experts bearing similar institutional status, but to select among a broader range of claimants to objective expertise.

---

[1] This example is based on a viral conspiracy theory alleging the non-existence of Finland. However, in contrast to many other conspiracy theories, it is unclear that anyone sincerely endorses this one.

Let us turn to some proposed strategies for selecting well, several of which were offered in a pioneering article by Goldman (2001). The first strategy discussed by Goldman is the direct evaluation of the content of experts' arguments. For present purposes, I take this strategy to be a non-starter. We are chiefly interested in those cases in which a layperson is not practically capable of reliably evaluating the cases for and against various claims. The inability to do so is, after all, the reason for which the layperson must identify the right expert(s) to which to defer.

Even if laypersons cannot meaningfully assess the *content* of experts' arguments, they might be better positioned to assess the *delivery* of experts' arguments. For example, Goldman suggests that laypersons might reasonably treat dialectical superiority—understood as effectiveness in responding to opponents' arguments—as an indicator of relative expertise (2001: 95). Notably, effectiveness in responding here cannot be understood to require responding with sound arguments, as the layperson is not positioned to evaluate this property of arguments.

According to Goldman, dialectical superiority is only a non-conclusive indicator of relative expertise. Unfortunately, even this diagnosis may be too optimistic. First, it is only rarely that laypersons have the opportunity to assess the relative dialectical skills of supposed experts, at least if such skills are interpreted narrowly as manifesting in debate contexts. Such debates occur relatively infrequently, in part for reasons discussed below. What is more, in debate contexts, dialectical superiority is often arguably an indicator that one's arguments are *unsound*. To see this, imagine a debate between a representative of the flat Earth conspiracy theory and a representative of the scientific orthodoxy. For the flat Earther, preparation for such a debate is a relatively easy matter. There are a limited number of widely-cited scientific grounds for accepting that the Earth is roughly spherical. The flat Earther need not develop sound counterarguments to these—only arguments that will likely appear plausible to laypersons. In contrast, the orthodox scientist's job is much more difficult. Because there is no sound basis on which the flatness of the Earth might be defended, the flat Earther can only rely on unsound arguments. There is virtually no limit to the number of unsound arguments that might be offered in favor of the flat Earth view. Admittedly, as an anonymous referee has helpfully pointed out, proponents of the flat Earth theory and other conspiracy theories tend to have favored arguments for their views. It might thus be thought that dialectical superiority can be demonstrated by preparing responses to such widely-cited arguments. The challenge for this approach, however, is that any such response is likely to be met by some further (likely ad-hoc) rejoinder that need not be constrained by the facts and thus for which it will be difficult to prepare. The orthodox scientist is thus likely to be surprised by certain arguments and may struggle to address these in a timely manner. Thus, while we might hope that legitimate experts would outperform pretenders in the context of debate, there is reason to expect that, even and perhaps especially when the pretender represents an especially absurd view, he or she will have the upper hand.

One might also hope to assess relative expertise by consideration of dialectical skills in a broader sense, where these may be manifested outside of the context of

structured debate. For example, one might think that exchanges between putative experts on social media or in blogs afford the opportunity to compare the ease and comprehensiveness with which the parties address the arguments put forward by their counterparts. One challenge for this suggestion is that, in contrast to watching a single debate, following the twists and turns in such a dialectic will require significant investment on the part of laypersons. Another is that, in such contexts, generating fake displays of dialectical skills that nonetheless impress laypersons will be relatively easy. Inauthentic experts may generate what they present as decisive refutations of their opponents' positions, and indeed might use large language models and other external resources to generate what appear to be credible arguments. By hypothesis, laypersons cannot assess the quality of these arguments directly and so, without access to information about the ways in which parties to a dialectic achieve dialectical superiority—or some counterfeit form of this—laypersons will likely be led astray.

Consider next some social strategies by which Goldman suggests expertise might be assessed. The first is by directly assessing the degree of putative expert support for various positions. Other things being equal, concurrence with the judgments of other putative experts speaks in favor of a given person's claim to expertise, and to the reliability of that person's claims. Second, one might consider credentials—including degrees, awards, publication histories, and professional roles—as a measure of others' attitudes toward a given supposed expert. Roughly, achievement of various credentials indicates others' confidence in a given person's claims to expertise.

The effectiveness of these social strategies will depend on various factors. First, appeal to further putative expert judgments will be sound only insofar as those putative experts are themselves reliable. Moreover, as Goldman notes, a lack of independence among putative experts significantly reduces the added value of the appeal to numbers.[2] Additionally, there are good reasons to doubt the reliability of some credentials as indicators of expertise. As Neil Levy (2022a: chap. 5) emphasizes, there exist a broad range of phony credentialing institutions that effectively compromise the significance of credentials. While experts themselves may be able to reliably distinguish between legitimate and illegitimate credentialing institutions, laypersons will likely find this far more challenging.

A further strategy suggested by Goldman is to consider putative experts' likely interests and biases. Other things being equal, the judgments of a putative expert whose interests align with the advancement of a claim carry less weight than those of a putative expert with little or no interest in the matter. Likewise, indications of bias on the part of putative experts indicate that their judgments should be given restricted significance. The strategy of assessing putative expert interests and biases has the advantage of being relatively tractable for non-experts. For example, whereas climate data and arguments based on such data are likely to be obscure from the perspective of the layperson, the motivation that those with connections to the fossil fuel industry have for pushing climate change denialism is easy to

---

[2] Notably, in real-world cases, there is good reason to expect putative expert judgments to be at least partially independent.

grasp. Nonetheless, this strategy is complicated by the existence of conspiracy theories that distort lay perceptions of the interests of putative authorities. For example, critics of the expert consensus can and do claim that proponents of the consensus are driven by perverse financial incentives.

The final strategy Goldman suggests is to consider the track records of putative experts. As Goldman notes, laypersons are often positioned to assess the accuracy of claims in given expert domains in the long run, even if they are not positioned to assess those claims when they are initially made. This is especially true when those claims take the form of predictions, the bases for which may be obscure, but the (non)realizations of which are readily observable.

The effectiveness of this latter strategy is heavily domain-dependent. Whereas those in the business of short-term predictions—meteorologists, for example—can quickly prove themselves through the success of their predictions, other domains afford considerably fewer opportunities for predictions assessable by laypersons. Some domains—history being a case in point—may simply admit of fewer predictions. Others, like climate science, often make predictions that are highly complex, assessable only in the long term, or some combination of these. Thus, the track records of certain experts may be difficult or impossible to assess from the lay perspective.

David Matheson (2005) builds on Goldman's discussion of dialectical superiority with two further considerations accessible to laypersons. One such consideration is the ability of putative experts to "manage evidence"—that is, to:

> [R]elate distinct pieces of ostensive evidence both to each other and to the question concerning which they are relevant; or, alternatively, an ability to see how the various pieces of ostensive evidence in one's possession fit together into one coherent whole, relevant to the question at hand. (D. Matheson 2005: 154)

This ability, which we might call *explanatory superiority*, reflects an underlying understanding of the relevant subject matter and the connections between various pieces of evidence pertaining to the issue.

One concern for the appeal to explanatory superiority is that some putative experts may exploit spurious connections to create the false appearance of strong evidence management. Consider, for example, the conspiracy theorist that weaves together a grand but highly misleading narrative to account for the available evidence and its significance. Notably, epistemologists studying conspiracy theories have drawn attention to the prominent role that errant data—data either unexplained by or conflicting with official accounts—play in conspiracy theories (Keeley 1999). In short, conspiracy theorists often draw attention to small anomalies to challenge official accounts and to support alternatives. For example, those promoting the conspiracy theory that the moon landing was a hoax point to the fact that, in the famous photo depicting the American flag on the surface of the moon, the flag appears to flutter despite the absence of wind on the surface of the moon. More generally, conspiracy theories often involve connecting the dots—that is, drawing connections between unrelated facts and events—a point

reflected in the tendency of conspiracy theorists to locate illusory patterns in noise (van Prooijen, Douglas, and De Inocencio 2018). The risk, then, is that appeals to explanatory superiority—or, more accurately, what is perceived by laypersons to be explanatory superiority—may systematically favor conspiracy theorists over their rivals.

Let us turn to a final strategy, also suggested by Matheson (2005). Matheson suggests that laypersons may assess the reliability of putative experts by, in part, attending to their willingness to seriously consider opposing positions and recalcitrant evidence and to own up to past mistakes. Roughly, then, Matheson suggests that laypersons ought to attend to the open-mindedness and humility of putative experts in order to facilitate judgments of reliability. Matheson speaks of this suggestion in terms of evaluations of the *moral superiority* of putative experts. Because open-mindedness and humility have been extensively discussed as intellectual virtues, I will discuss this strategy in terms of *intellectual virtue superiority*.

This final suggestion from Matheson has some immediate appeal. First, it is plausible that laypersons can and do attend to displays of intellectual character, and sometimes treat these as indicators of reliability. Second, the strategy appears sound on its face. If putative experts are not willing to seriously engage with alternatives, it seems that they would be unlikely to correct false beliefs. Similarly, if experts do not own up to past mistakes, they are less likely to learn from those past mistakes. Thus, the intellectual virtues of open-mindedness and humility would seem to facilitate reliability. Moreover, the possession or lack of such virtues seems on the face of things to be accessible to laypersons. One need not be an expert to see that a putative expert is highly dismissive of alternatives or fails to acknowledge past mistakes. The strategy of considering putative experts' intellectual virtues thus seems very promising on its face, at least as a supplement to the other strategies considered above. However, beginning in the next section, I argue that the attempt to assess expertise by attending to displays of intellectual virtue is likely to lead laypersons astray. I then argue that this problem is not easily solved.

To conclude this section, it is worth emphasizing that the strategies discussed, including attention to intellectual virtue superiority, need not be construed simply as strategies for making one-to-one comparisons between individual putative experts. In addition to comparing individuals, these strategies might be used to make comparisons between communities. For example, a layperson might assign greater credence to explanations offered by large communities that, collectively, exhibit dialectical and explanatory superiority relative to rival communities. Moreover, insofar as collectives can themselves exhibit intellectual virtues and vices, as discussed by Miranda Fricker (2010), Alessandra Tanesini (2023), and myself (Harris 2021), it is possible in principle for laypersons to assign credibility to collectives according to the intellectual characters of these communities. More generally, the challenge of identifying experts need not be construed as a challenge of sorting between individuals, but can also be understood as the challenge of determining how to assess the relative credibility of collectives.

2. Intellectual Virtue Signaling

Beginning in this section, I argue that attempts to locate experts by using displays of intellectual virtue are likely to go awry. The basic concern resembles many of those raised for the strategies discussed in the previous section. There, I noted that many techniques for identifying experts can misleadingly suggest the reliability of non-experts, and indeed of the especially unreliable, while calling into doubt the reliability of legitimate experts. For example, dialectical superiority may favor those that are not constrained by the truth. Similarly, attending to displays of intellectual virtue as signs of expertise is likely to lead laypersons astray. This is for two interconnected reasons.

The first and more basic concern is that unreliable persons often put on ostentatious displays of what appear to be, but need not actually be, intellectual virtues. Levy (2023) has coined the term *intellectual virtue signaling* to capture this practice. Levy understands intellectual virtue signaling as an analog of *moral virtue signaling*, which is defined as "moral talk ostensibly about some state of affairs or action that is designed in part to draw attention to the person doing the signaling" (Levy 2023: 312). In short, moral virtue signalers aim not—or not exclusively—to *do* good, but to *seem* good. Likewise, intellectual virtue signalers may or may not actually possess intellectual virtues but aim at least in part to seem as if they do.

It may be objected that this definition is excessively broad, as being *partly* motivated to seem good is extremely common and thus the present definition would overdiagnose instances of behavior as virtue signaling[3] (cf. Tosi and Warmke 2020: 20). In their work on moral virtue signaling—*moral grandstanding*, in their terminology—Justin Tosi and Brandon Warmke (2020) offer a useful test meant to indicate more clearly when an act involves moral grandstanding. On their view, an act counts as a case of moral grandstanding only if[4] the actor would be disappointed to learn that the audience was not impressed by the moral qualities displayed by the act (Hill and Fanciullo 2023: 4; Tosi and Warmke 2020: 19). I have some reservations about using proneness to disappointment as a measure of the strength of one's desires. When one enters a lottery, one may very strongly desire to win, even though one will not be disappointed if one does not win. This is because, while one wants to win, one expects not to win. More generally, disappointment seems to depend on both the strength of one's desires and on one's expectations. Still, I think Tosi and Warmke's test offers at least a good heuristic for distinguishing between cases that do and do not involve virtue signaling. If one finds Tosi and Warmke's test inadequate, alternative ways of distinguishing between moral grandstanding (or virtue signaling) and cases in which one is merely partly motivated by concerns about one's reputation, without thereby virtue signaling, are available. Evan

---

[3] Thus, for example, Jesse Hill and James Fanciullo describe the "whole-hearted grandstander" as one who genuinely believes in a cause but is also significantly motivated by a desire for recognition as virtuous. Thanks to an anonymous referee for raising this objection.

[4] Elsewhere, Tosi and Warmke (2016: 202) indicate that the desire for recognition is not a necessary condition on moral grandstanding, but is typical of paradigmatic cases of moral grandstanding.

Westra (2021), for example, writes that virtue signaling occurs just when one's motivation to contribute to public moral discourse is only secondary to one's desire for recognition. We might understand this to mean that, if the agent's motivations were at odds then, other things being equal, the desire for recognition would win out. This would be a relatively strong condition on virtue signaling. Alternatively, one might say that an act amounts to virtue signaling if one's desire for recognition is a difference-maker, in the sense that one would not have performed the act if one lacked the desire for recognition. Because my aim here is not to analyze virtue signaling, I will remain neutral between these approaches.

Catharine Saint-Croix (forthcoming) discusses intellectual virtue signaling in terms of *epistemic virtue signaling*. As Saint-Croix emphasizes, intellectual virtue signaling may, but need not, consist in an attempt to display a particular virtue. In some cases, one might aim to present oneself as having a particular intellectual virtue. For example, one might boast of one's consumption of newspapers representing a range of ideological perspectives as a way of signaling one's open-mindedness. However, intellectual virtue signaling might also consist in attempts to convey a general impression of intellectual superiority, independent of any particular intellectual virtue. Consider, for example, the smug use of the phrase "I don't own a television." Because I am concerned here with the prospects of using displays of specific intellectual virtue as indicators of reliability, I focus on cases in which individuals or collectives aim to present themselves as having specific intellectual virtues.

With this point in mind, let us consider how the above approaches might be applied to derive different accounts of intellectual virtue signaling in particular. On one possible view, one engages in virtue signaling only if one does so with a desire to be recognized as possessing intellectual virtue that is sufficiently strong that one will be disappointed if this desire is not satisfied. As the discussion above anticipates, this approach is complicated by the fact that proneness to disappointment plausibly depends on both the strength of desire and one's expectations. One might strongly hope to impress an audience with one's display of intellectual virtue but, because one anticipates a tough crowd, not expect to impress the audience in this way. In such a case, it seems plausible that one might be engaging in intellectual virtue signaling even though one would not be disappointed by failure to achieve recognition as intellectually virtuous. On an alternative approach, one's desire to be recognized as intellectually virtuous must be stronger than the other motivating desires behind one's actions.[5] For example, if one posts a long social media thread about a foreign conflict and is motivated in part by the desire to inform one's followers, but desires even more strongly to be recognized as intellectually virtuous by those followers, then, on this second approach, one engages in intellectual virtue signaling. Finally, to illustrate the third approach, we might imagine some individual who has some motivation to inform her followers by way of a social media thread, but whose desire to do so is not sufficiently strong, by itself, to bring her to do so. On the third approach, her

---

[5] Saint-Croix (forthcoming) explicitly rejects the view that intellectual virtue signaling only occurs when one's desire for recognition is dominant.

posting of the threat will constitute intellectual virtue signaling if she posts it only because her desire to inform is supplemented by a desire for recognition as an intellectually virtuous person—that is, if her latter desire is a difference-maker with respect to whether she performs the action in question.

My aim here is not to defend any particular account of virtue signaling, and is instead to argue that the practice of virtue signaling complicates one proposed strategy for identifying experts. Thus, I will not attempt a decisive argument for any of these approaches here. However, it is worth noting that, with the aim of assessing the effects of intellectual virtue signaling in mind, there is some reason to favor a relatively demanding account of the desire for recognition. If we suppose that acts may constitute intellectual virtue signaling in virtue of just any attendant desire for recognition as intellectually virtuous, then intellectual virtue signaling is rampant—indeed I confess that I might be engaging in some right now (cf. Levy 2023, 320–21)! Thus, it may be preferable to favor a relatively demanding desire for recognition condition so that the resultant definition better accords with instances that would be recognizable as intellectual virtue signaling. In any case, I will focus below on cases in which the desire for recognition as virtuous is especially strong. The thesis of this article might thus be framed in terms of the effects of intellectual virtue signaling, or a particularly egregious subclass of it, depending on one's favored approach to the desire for recognition condition.

Levy's core example of intellectual virtue signaling is the practice, familiar on social media, of non-experts offering lengthy and typically heterodox takes on a wide variety of hot topics. Thus, for example, the same person might offer what seem to be nuanced and skeptical takes on events ranging from domestic economic matters, foreign wars, and global pandemics. In so doing, that person puts on a display of apparent intelligence, intellectual courage, and autonomy (Levy 2023, 316). Insofar as this display is effective, the intellectual virtue signaler is likely to draw attention away from legitimate experts, to cause confusion about who possesses legitimate expertise, and, in this way, to generate broader confusion about the truth (Levy 2023, 320). As Levy emphasizes, the latter problem arises in part because the intellectual virtue signaler is incentivized to offer heterodox judgments as a means of maximizing the attention they receive. Especially on social media, one's takes must be *hot* to garner attention. Here, I focus on how intellectual virtue signaling can complicate the process of identifying experts according to the strategy Matheson (2005) suggests, and the extent to which this challenge can be solved.

Levy focuses on the signaling of intelligence, courage, and autonomy. Let us turn to a further virtue that has received considerable philosophical attention in recent years (Battaly 2018b; 2018a): the virtue of open-mindedness (Bland 2022; Levy 2022b; Riggs 2010). As a first pass, to be open-minded is to be willing to take seriously alternatives to one's own views. As we will see below, this initial gloss on open-mindedness is overly simple, but it will suffice for now. I focus here on open-mindedness because the exploitative signaling of this virtue appears to be particularly prevalent. Thus, for example, it is commonplace for conspiracy theorists to signal their possession of this virtue by refusing to take for granted official accounts of events, by attending to a broad range of possibilities, by

challenging representatives of the orthodoxy to debates, and so on. Thus, for example, while many would refuse to take seriously the possibility that the September 11 attacks were a false flag, conspiracy theorists proclaiming a commitment to open-mindedness might well treat this possibility as worth investigating.

Notably, while conspiracy theorists outwardly display a commitment to open-mindedness, it is not clear that these displays are indicative of their actual practices. For example, some empirical evidence indicates that open-mindedness is negatively correlated with belief in conspiracy theories (Swami et al. 2014). While the empirical evidence is far from conclusive, it is at least not obvious that conspiracy theorists' commitment to open-mindedness is genuine. This is especially true of *promoters* of conspiracy theories—those that broadcast and often profit from broadcasting conspiracy theories—as opposed to those merely taken in by them. At a minimum, we might say that displays of open-mindedness are signals of reliability too easily feigned to be useful in discerning genuine experts.

The same is true of other displays of intellectual virtue, for example the virtue of intellectual autonomy. On the face of it, intellectual autonomy is the tendency to think for oneself. Elizabeth Fricker, for example, writes that the intellectually autonomous person "takes no one else's word for anything, but accepts only what she has found out for herself, relying only on her own cognitive faculties and investigative and inferential powers" (E. Fricker 2006). As with open-mindedness, this initial gloss on intellectual autonomy is arguably too simple and will be complicated below. Intellectual autonomy, so construed, is plausibly the chief virtue of conspiracy theorists, and "think for yourself," "do your own research," and related slogans are regularly repeated by conspiracy theorists (Ballantyne, Celniker, and Dunning 2022; Buzzell and Rini 2023; Levy 2022c).

Notably, promoters of conspiracy theories do not simply display the virtue of epistemic autonomy in their own practice. Rather, conspiracy theorists counsel that others ought to think for themselves. Such advice indicates the absence of an agenda, and therefore encourages trust in conspiracy theory promoters. As with open-mindedness, there is little reason to think that conspiracy theorists are *actually* more epistemically autonomous than others, or that exhortations to think for oneself represent a sincere commitment to the promotion of epistemic autonomy. As to the former point, empirical research indicates that conspiracy theorists treat others' judgments as evidence in much the same way that non-conspiracy theorists do (Altay et al. 2023). As to the latter, the exhortation to think for oneself is regularly accompanied by the presentation of misleading evidence favoring foregone conspiratorial conclusions.

Thus far in this section, I have argued that the strategy of identifying experts by attending to displays of intellectual virtue is easily exploitable by those that fake displays of intellectual virtue. My core examples have been conspiracy theorists and, in particular, conspiracy theory promoters. However, the point generalizes to a range of other non-experts. Proponents of young earth creationism, for example, regularly challenge orthodox evolutionary biologists to debates—in displays of apparent open-mindedness. Similarly, promoters of alternative medical practices often bemoan that their positions are not given a fair hearing alongside those of

orthodox medicine and encourage individuals to explore alternative medicine for themselves.

These examples hint at a second concern for attention to displays of intellectual virtue as indicators of expertise. In short, legitimate experts often fail to display intellectual virtues in ways corresponding to displays by their counterparts and, indeed, often seem to display intellectual vices. Proponents of evolution by natural selection often refuse to engage with creationists, and even discourage others from doing so (Etchells 2014). Similarly, many mainstream scientists and other epistemic authorities refuse to publicly engage with conspiracy theories and their promoters (Lee 2023), and recommend that others do the same. Indeed, it has been argued that the failure of experts to engage seriously with apparent evidence against their views contributes to the creation and spread of conspiracy theories (Brooks forthcoming). Moreover, legitimate experts often seem to shun epistemic autonomy, both in themselves and others. If the conspiracist mantra is "think for yourself" that of many experts is to "trust the science."

Insofar as a layperson uses displays of apparent intellectual virtue as cues of reliability, that person is likely to be led systematically astray. This is not only because non-experts regularly engage in ostentatious displays of apparent intellectual virtue, but also because legitimate experts often seem to behave non-virtuously and, indeed, intellectually viciously. In fact, certain practices—for example conspiracists' and creationists' challenging of scientists to debates—serve simultaneously to signal virtue and to seemingly expose the vices of their counterparts. It might, however, seem that the problem identified here is easily solved, at least in part. All it would take is for legitimate experts to more deliberately display intellectual virtue. They might, for example, go to greater lengths to engage in debates with representatives of alternative views, thereby displaying open-mindedness. Similarly, they might display a respect for epistemic autonomy by eschewing reliance on other putative experts and by encouraging laypersons to consider, explore and weigh the evidence for themselves. In the next section, however, I argue that such shifts in practice would often do more harm than good.

3. The Limits of Virtue

On the face of things, the suggestion that genuine experts can address the problem described above by displaying greater intellectual virtue seems like win-win. Greater open-mindedness, humility, and intellectual autonomy on the part of experts would make such experts easier to identify, while also improving the quality of their inquiries. Or so one might think. However, this suggestion is overly optimistic. As I argue in this section, greater displays of intellectual virtue would either be insincere and counterproductive or would *reduce* the ability of experts to provide genuine valuable information.

Suppose, first, that experts merely *pretend* to be intellectually virtuous—that is, that they inauthentically signal intellectual virtue. One way they might do so is by emulating those non-experts that signal commitment to epistemic autonomy. Experts might, for example, encourage laypersons to do their own research. To

facilitate this, experts might release more data without offering their own interpretations of the data. While such a shift might encourage greater trust in genuine experts, any such victory would be pyrrhic. Unless experts offer their own interpretations, increased trust in actual experts would do little to improve the epistemic situations of laypersons. Moreover, expert exhortations to do one's own research would likely do more harm than good. This is because, as a general rule, "do your own research" is a bad epistemic policy for laypersons to follow (Huemer 2005). Laypersons are distinguished, in large part, by their lack of expertise—which is likely to include the inability to reliably interpret relevant data in expert domains. Admittedly, some laypersons *could* develop the needed expertise to reliably interpret data in certain expert domains. However, even in this case, they would remain non-experts in other domains. Moreover, insofar as the development of expertise enables former laypersons to reach the same judgments already reached by experts, this process would involve a redundancy of expertise.

Suppose, alternatively, that experts work to *actually* embody the sorts of virtues described above, including open-mindedness, humility, and epistemic autonomy. That is, suppose that experts themselves opt to be more open-minded, humbler, and more epistemically autonomous. A given scientific expert might, for example, display greater autonomy by refusing to take for granted the findings of other scientists, and opting instead to, as the conspiracy theorist exhorts, do her own research. In so doing, the scientist might display her intellectual virtue to laypersons, thereby showing herself to be reliable. There are at least two severe problems with this suggestion. The first is practical. A scientist who proposed to repeat research already conducted by others would struggle to attract funding for doing so. The latter, and perhaps more serious concern, is that such a scientist would be unable to get far without relying on the results of others. As epistemologists have emphasized, we are pervasively epistemically dependent on others (Hardwig 1985), such that laypersons and experts alike owe much of their knowledge to others. In the context of science, in particular, a refusal to depend on others would serve only to stunt the progress of science, as effort would be wasted on attempts to individually do what has already been done by others.[6]

Consider, next, the virtues of humility and open-mindedness. An expert might demonstrate a commitment to such virtues by, for example, engaging in debates with those holding heterodox views, by hearing out proponents of those views, and by testing a broad range of alternatives. For example, during the COVID-19 pandemic, medical experts might have demonstrated these virtues by debating proposed alternative treatments, giving their attention to those promoting such alternatives, and indeed putting proposed alternatives to the test. In so doing, they

---

[6] None of this is to say that scientists should *never* doublecheck the work of others. The credibility of some sciences has been damaged by questionable practices like p-hacking and data dredging (Anvari and Lakens 2018; Smith 2023; Wingen, Berkessel, and Englich 2020), and the replication crisis that is attributable in part such practices. Attempts at replication can promote integrity on the part of other scientists and thereby promote trust in scientific findings.

might have attracted the trust of those that, like some critics of mainstream views about the pandemic, claimed that proper scientists are willing to question everything.

One problem in this case is that doing such things would be, by and large, a waste of time. A second problem is that, if experts strive to be genuinely open-minded by, for example, giving equal hearing to ideas long thought to be non-credible by the broader community, some are likely to be led astray. At best, they will waste their time re-learning what they have already learned from others. At worst, they will themselves form mistaken judgments. The broader problem is that there is little reason to expect that the qualities that laypersons think promote good science actually do so. The traits that would make one a trusted science communicator are not necessarily the same traits that would make one a good scientist. Because scientific progress depends in large part on the division of labor, scientific success is often promoted by dependency, closed-mindedness, arrogance, and other traits that, while perhaps damaging at the individual level, promote the exploration and thorough consideration of a broad but limited range of reasonable alternatives at the collective level. The preceding point resembles in broad strokes the point, made by a variety of authors (Hull 1997; Kitcher 1995; Smart 2018), that successful science does not depend on individual scientists being motivated to contribute to the development of collective knowledge. Instead, as has been said of markets, some suggest that the structure of science harnesses private vice into collective benefits[7]—in this case shared knowledge. Rather than focusing broadly on the relation between the motives of scientists and the success of science, I focus here instead on how the lack of specific virtues among individuals and collectives may serve the collective epistemic good among practitioners in expert domains.

It may be objected that the apparent tension between epistemic virtue on the part of experts and collective success within expert domains that I have emphasized here is contingent on an overly general and perhaps naïve conception of epistemic virtue. For example, one might say that open-mindedness is not, as I have defined it above, best captured simply in terms of a willingness to take seriously alternatives to one's views. The open-minded person need not take seriously just any preposterous view that happens to come to his or her attention. Rather, the open-minded person need only be open to a relatively narrow set of possibilities that meet some threshold of plausibility. Similarly, one might say that the virtue of autonomy does not require its possessor to eschew reliance on others. Instead, one might say, the virtue of intellectual autonomy is consistent with reliance on others, so long as that reliance is reflective of one's own good judgments as to who to rely upon, and how and when to rely upon them (J. Matheson 2021, 182–83). Alternatively, one might say that the mistake is to think of epistemic autonomy as a virtue at all, at least when this virtue is not regulated by the additional virtue of epistemic interdependence (Battaly 2021).

These alternatives raise interesting questions as to what makes a virtue a virtue. Perhaps, one might say, the cases highlighted here suggest that whether a given trait constitutes a virtue is dependent on its effects. This view, which is endorsed

---

[7] Taking inspiration from Bernard Mandeville's *Fable of the Bees*, Paul R. Smart (2018) refers to this phenomenon as *Mandevillian intelligence*.

by some virtue responsibilists (Zagzebski 1996, pt. II), would account for why there are limits to the alternatives that an open-minded person would countenance. It would also account for why autonomy—if understood to involve non-reliance on others—is no virtue.

I do not aim to determine here whether this approach to the epistemic virtues is to be preferred to a simpler one. Instead, what I want to highlight here is that—even if epistemic virtues are best construed along the relatively complex lines suggested here —there is little reason to expect ordinary members of the public to have such a sophisticated account of the virtues. What individuals do to signal their own intellectual virtue, and what acts individuals interpret as indicative of intellectual virtue, will depend on the norms operative in their specific social epistemic contexts (cf. Saint-Croix forthcoming). Suppose, for example, that one puts on a display of not believing anything based on the word of others. Even if this is not a genuine indicator of epistemic autonomy, and indeed even if intellectual autonomy is not a virtue at all,[8] such an act might be undertaken in the hopes of indicating one's intellectual virtues, and might be interpreted by others as doing so. Thus, even if we suppose that deference to those who *actually* possess intellectual virtues would be a reliable belief-forming strategy, mistaken conceptions of what the intellectual virtues involve, and indeed what the virtues are, will likely lead laypersons to trust those who lack genuine intellectual virtues. By a similar token, misunderstandings about intellectual virtue may lead laypersons to distrust those who exhibit genuine intellectual virtues. Even if there is in fact no tension between virtue on the part of experts and success in expert domains, individuals untutored in virtue epistemology are likely to perceive such a tension.

Let us illustrate the preceding points with an example. Given the assumption that members of the public treat displays of apparent virtue and vice as a basis on which to apportion trust, the mere appearance of such a tension is sufficient to result in *mis*allocations of trust. If the conspiracist podcaster displays traits that ordinary persons perceive to be illustrative of virtue, while the legitimate expert displays traits that ordinary persons perceive to be illustrative of vice, then, even if these perceptions rest on a misunderstanding of the epistemic virtues, the public will place too much trust in the podcaster and too little trust in the legitimate expert. More generally, when laypersons use intellectual virtue as a marker of expertise, the best they can do is identify those that *seem, from the lay perspective*, to be virtuous. As I have argued, the attempt to identify legitimate experts using such cues is likely to lead laypersons astray.

4. Concluding Remarks

Laypersons cannot be expected to competently form judgments about matters in expert domains without guidance by legitimate experts. Thus, laypersons require strategies for identifying legitimate experts. Any such strategy must be both reliable and accessible for individuals lacking in domain knowledge. On the face

---

[8] Thanks to an anonymous referee for raising this point.

of things, the identification of legitimate experts through assessment of cognitive character seems like a promising strategy insofar as it is both feasible and continuous with familiar practices of character assessment. However, as I have argued here, such a strategy is likely to go systematically awry—as displays of apparent intellectual virtue are not reliably correlated with—and may indeed be negatively correlated with—actual expertise. One might, on this basis, draw one of at least two different conclusions.

The first, perhaps more pessimistic conclusion, is that attention to displays of intellectual virtue is, at best, useless as a strategy for identifying legitimate experts. One tempted by this conclusion might argue that laypersons would be better served by learning about how, for example, the structure of science can allow science to succeed in spite of, and indeed perhaps because of, the vices of scientists. In this way, laypersons might be brought to trust a consensus reached by scientists, even if they do not regard individual scientists as worthy of trust.

A more optimistic conclusion is that, if the identification of experts through attention to the virtues and vices of putative experts is to succeed, it will require laypersons to adopt a more sophisticated approach to the virtues along the lines of the one sketched in section three. Laypersons must learn, perhaps by being taught, that pretenders to expertise sometimes look fair, while feeling foul. While ambitious, it is perhaps not unrealistic that laypersons might learn this lesson. Certain outward displays look like signs of friendliness, but most of us have learned that such displays do not reliably indicate the good character of used car salesmen and the like. Likewise, laypersons might learn that ostentatious displays of willingness to debate, to consider alternatives, to take nothing for granted, to think for oneself, and so on, need not indicate the possession of epistemic intellectual virtues, properly understood. Through the development of more sophisticated understandings of the intellectual virtues and their indicators, the viability of this strategy for identifying experts might be increased.

These conclusions need not be regarded as entirely incompatible alternatives. Given that even experts are fallible and otherwise flawed, reliance on the collective and institutional virtues of expert domains is often a surer strategy than trust in particular individuals. At the same time, given our social natures, it would likely be unrealistic to expect laypersons to eschew relations of personalized trust toward putative experts altogether. Consequently, the best path forward for laypersons searching for truth in expert domains is likely a combination of defeasible trust in persons that display genuine intellectual virtues, and reliance on the collective and institutional virtues of expert domains to constrain and harness the negative inclinations of individuals.

KEITH RAYMOND HARRIS ⓘD
UNIVERSITY OF VIENNA
keithraymondharris@gmail.com

References

Altay, Sacha, Kenzo Nera, Waqas Ejaz, Céline Schöpfer, and Frédéric Tomas. 2023. "Conspiracy Believers Claim to Be Free Thinkers but (Under)Use Advice like Everyone Else." *British Journal of Social Psychology*, May, bjso.12655. https://doi.org/10.1111/bjso.12655.

Anvari, Farid, and Daniël Lakens. 2018. "The Replicability Crisis and Public Trust in Psychological Science." *Comprehensive Results in Social Psychology* 3 (3): 266–86. https://doi.org/10.1080/23743603.2019.1684822.

Ballantyne, Nathan, Jared B. Celniker, and David Dunning. 2022. "'Do Your Own Research.'" *Social Epistemology* 0 (0): 1–16. https://doi.org/10.1080/02691728.2022.2146469.

Battaly, Heather. 2018a. "Closed-Mindedness and Dogmatism." *Episteme* 15 (3): 261–82. https://doi.org/10.1017/epi.2018.22.

Battaly, Heather. 2018b. "Can Closed-Mindedness Be an Intellectual Virtue?" *Royal Institute of Philosophy Supplements* 84 (November): 23–45. https://doi.org/10.1017/S135824611800053X.

Battaly, Heather. 2021. "Intellectual Autonomy and Intellectual Interdependence." In *Epistemic Autonomy*, by Jonathan Matheson and Kirk Lougheed, 1st ed., 153–72. New York: Routledge. https://doi.org/10.4324/9781003003465-12.

Bland, Steven. 2022. "Interactionism, Debiasing, and the Division of Epistemic Labour." In *Social Virtue Epistemology*, by Mark Alfano, Colin Klein, and Jeroen De Ridder, 1st ed., 15–38. New York: Routledge. https://doi.org/10.4324/9780367808952-3.

Brooks, Patrick. forthcoming. "On the Origin of Conspiracy Theories." *Philosophical Studies*. https://philarchive.org/rec/BROOTO-6.

Buzzell, Andrew, and Regina Rini. 2023. "Doing Your Own Research and Other Impossible Acts of Epistemic Superheroism." *Philosophical Psychology* 36 (5): 906–30. https://doi.org/10.1080/09515089.2022.2138019.

Etchells, Pete. 2014. "Bill Nye v Ken Ham: Should Scientists Bother to Debate Creationism?" *The Guardian*, February 5, 2014, sec. Science. https://www.theguardian.com/science/head-quarters/2014/feb/05/bill-nye-vs-ken-ham-creationism-science-debate.

Fricker, Elizabeth. 2006. "Testimony and Epistemic Autonomy." In *The Epistemology of Testimony*, edited by Jennifer Lackey and Ernest Sosa, 1st ed., 225–50. Oxford University Press: Oxford. https://doi.org/10.1093/acprof:oso/9780199276011.003.0011.

Fricker, Miranda. 2010. "Can There Be Institutional Virtues?" *Oxford Studies in Epistemology* 3: 235–52.

Goldman, Alvin I. 2001. "Experts: Which Ones Should You Trust?" *Philosophy and Phenomenological Research* 63 (1): 85–110. https://doi.org/10.1111/j.1933-1592.2001.tb00093.x.

Hardwig, John. 1985. "Epistemic Dependence." *The Journal of Philosophy* 82 (7): 335. https://doi.org/10.2307/2026523.

Harris, Keith Raymond. 2021. "Collective Intellectual Humility and Arrogance." *Synthese* 199 (3): 6967–79. https://doi.org/10.1007/s11229-021-03101-y.

Hill, Jesse, and James Fanciullo. 2023. "What's Wrong with Virtue Signaling?" *Synthese* 201 (4): 117. https://doi.org/10.1007/s11229-023-04131-4.

Huemer, Michael. 2005. "Is Critical Thinking Epistemically Responsible?" *Metaphilosophy* 36 (4): 522–31.

Keeley, Brian L. 1999. "Of Conspiracy Theories." *The Journal of Philosophy* 96 (3): 109. https://doi.org/10.2307/2564659.

Lee, Bruce Y. 2023. "Why Peter Hotez Should Not Debate RFK Jr. On The Joe Rogan Experience." Forbes. 2023. https://www.forbes.com/sites/brucelee/2023/06/26/why-peter-hotez-should-not-debate-rfk-jr-on-the-joe-rogan-experience/.

Levy, Neil. 2022a. *Bad Beliefs: Why They Happen to Good People*. First edition. New York: Oxford University Press.

Levy, Neil. 2022b. "Narrowing the Scope of Virtue Epistemology." In *Social Virtue Epistemology*, by Mark Alfano, Colin Klein, and Jeroen De Ridder, 1st ed., 113–30. New York: Routledge. https://doi.org/10.4324/9780367808952-15.

Levy, Neil. 2022c. "Do Your Own Research!" *Synthese* 200 (5): 356. https://doi.org/10.1007/s11229-022-03793-w.

Levy, Neil. 2023. "Intellectual Virtue Signaling." *American Philosophical Quarterly* 60 (3): 311–24. https://doi.org/10.5406/21521123.60.3.07.

Matheson, David. 2005. "Conflicting Experts and Dialectical Performance: Adjudication Heuristics for the Layperson." *Argumentation* 19 (2): 145–58. https://doi.org/10.1007/s10503-004-7041-0.

Matheson, Jonathan. 2021. "The Virtue of Epistemic Autonomy." In *Epistemic Autonomy*, 173–94. Routledge.

Prooijen, Jan-Willem van, Karen M. Douglas, and Clara De Inocencio. 2018. "Connecting the Dots: Illusory Pattern Perception Predicts Belief in Conspiracies and the Supernatural." *European Journal of Social Psychology* 48 (3): 320–35. https://doi.org/10.1002/ejsp.2331.

Riggs, Wayne. 2010. "Open-Mindedness." *Metaphilosophy* 41 (1–2): 172–88. https://doi.org/10.1111/j.1467-9973.2009.01625.x.

Saint-Croix, Catharine. forthcoming. "Epistemic Virtue Signaling and the Double Bind of Testimonial Injustice." *Philosophers' Imprint*.

Smart, Paul R. 2018. "Mandevillian Intelligence." *Synthese* 195 (9): 4169–4200. https://doi.org/10.1007/s11229-017-1414-z.

Smith, Gary. 2023. *Distrust: Big Data, Data-Torturing, and the Assault on Science*. Oxford: Oxford University Press.

Swami, Viren, Martin Voracek, Stefan Stieger, Ulrich S. Tran, and Adrian Furnham. 2014. "Analytic Thinking Reduces Belief in Conspiracy Theories." *Cognition* 133 (3): 572–85. https://doi.org/10.1016/j.cognition.2014.08.006.

Tanesini, Alessandra. 2023. "Intellectual Arrogance: Individual, Group-Based, and Corporate." *Synthese* 202 (1): 1–20. https://doi.org/10.1007/s11229-023-04224-0.

Tosi, Justin, and Brandon Warmke. 2016. "Moral Grandstanding." *Philosophy & Public Affairs* 44 (3): 197–217. https://doi.org/10.1111/papa.12075.

Tosi, Justin, and Brandon Warmke. 2020. *Grandstanding: The Use and Abuse of Moral Talk*. New York (N.Y.): Oxford University Press.

Westra, Evan. 2021. "Virtue Signaling and Moral Progress." *Philosophy & Public Affairs* 49 (2): 156–78. https://doi.org/10.1111/papa.12187.

Wingen, Tobias, Jana B. Berkessel, and Birte Englich. 2020. "No Replication, No Trust? How Low Replicability Influences Trust in Psychology." *Social Psychological and Personality Science* 11 (4): 454–63. https://doi.org/10.1177/1948550619877412.

Zagzebski, Linda Trinkaus. 1996. *Virtues of the Mind: An Inquiry Into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge University Press.