Authority or Autonomy?
Philosophical and Psychological Perspectives on Deference to Experts

Alex Worsnip
Department of Philosophy, University of North Carolina at Chapel Hill

Devin Lane
Department of Philosophy, University of North Carolina at Chapel Hill

Sam Pratt
Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill

M. Giulia Napolitano
School of Philosophy, Erasmus University Rotterdam

Kurt Gray
Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill

Jeffrey A. Greene
School of Education, University of North Carolina at Chapel Hill

Abstract. Several decades of work in both philosophy and psychology acutely highlights our limitations as individual inquirers. One way to recognize these limitations is to defer to experts: roughly, to form one's beliefs on the basis of expert testimony. Yet, as has become salient in the age of Brexit, Trumpist politics, and climate change denial, people are often mistrustful of experts, and unwilling to defer to them. It's a trope of highbrow public discourse that this unwillingness is a serious pathology. But to what extent is this trope accurate? Answering this requires us to settle both a normative question— under exactly what conditions *ought* we to defer to experts?—and an empirical question—under what conditions *are* people willing to defer to experts? The first question has been investigated primarily by philosophers; the second, primarily by psychologists. Yet there is little work integrating these literatures and putting together their results. The aim of this review article is to begin this task, enabling us to begin reaching conclusions about how much real practices of deference diverge from the ideal. We present an opinionated guide to relevant work from both philosophy and psychology, and note places where the literature has important gaps.

Keywords: deference to experts; epistemic autonomy; epistemic authority; trust in experts; philosophy of expertise; psychology of expertise

0. Introduction

There is a long and venerable intellectual tradition that stresses the value of intellectual autonomy— of figuring things out for oneself. Yet several decades of work in both philosophy and psychology, as

well as much of our everyday experience, acutely highlight our limitations as individual inquirers.[1] These limitations only increase in magnitude and importance in the context of our modern "infodemic" of (mis)information (World Health Organization, 2020), and as our social and intellectual world becomes increasingly complex and contested. The investigation of many modern controversies—for example, climate change, vaccine efficacy, or genetically modified organisms—requires background knowledge, skill, and resources that no one person has on their own. Very plausibly, to fail to recognize these limitations is to fail to evince appropriate intellectual humility—which, on a popular account, consists in "proper attentativeness to, and owning of, one's intellectual limitations" (Whitcomb et al. 2017: 520).

By contrast, one way to recognize these limitations—and thus, to display intellectual humility and to inquire more responsibly—is to appropriately defer to experts: that is, roughly, to form one's beliefs on the basis of expert testimony. Almost by definition, experts about some matter tend to possess more relevant knowledge and evidence than laypeople, and have greater capacity to correctly interpret it. When people are willing to defer to experts, rather than try to figure things out on their own, they (often) evince recognition of these advantages.

Yet, as has become especially salient in the age of Brexit, Trumpist politics, climate change denial, and more, people are often mistrustful of experts, and unwilling to defer to them. It's a trope of highbrow public discourse that this unwillingness to defer to experts is a serious pathology of contemporary society. But to what extent is this trope accurate? Answering that question requires us to settle at least two further questions, one normative and one empirical. The normative question is this: under exactly what conditions *ought* we to defer to experts? And the empirical question is this: under what conditions *are* people willing, or unwilling, to defer to experts? Only when we are in a position to answer both of these questions in detail can we see the extent to which laypeople's actual practices diverge from what is ideal.

Since the first question is normative, it has been investigated primarily by philosophers; since the second question is empirical, it has been investigated primarily by social scientists, especially psychologists. Yet to the best of our knowledge, there is little to no work attempting to systematically integrate these two literatures and put together their results. The aim of this review article is to begin the task of doing this, in the hope of enabling us to start to reach conclusions about the extent to which real practices of deference diverge from the ideal. We present a comparative guide to existing work from both philosophy and psychology relevant to our topic, and note places where the literature has important gaps yet to be filled. In general, where there is controversy in the literature (either normative or empirical), we have not attempted to take a definitive stance but to present the competing views and to summarize some of the more important considerations that have been offered in favor

[1] See the essays in Ballantyne & Dunning (2022) for a recent overview.

of them. Likewise, we acknowledge concerns about the replicability, universality, and theory crises in psychological research, and that these can call into question findings and inferences from some studies cited here (cf. Greene, 2022; Hutmacher & Franz, 2024). That said, we find the psychological evidence presented here useful and directionally informative, if not definitive.

We will structure our approach by considering a range of variables that may affect the (normative) propriety and/or the (empirical) prevalence of deference to experts. But first, we need some conceptual preliminaries; specifically, to understand what the various literatures have in mind by 'deference' and by 'experts.'

1. Clarifying the topic

a) *What is deference?*

Above, we roughly characterized deference to experts in terms of forming beliefs on the basis of expert testimony[2]—presumably, we can add, on the topic on which they are an expert. But this is vague, since someone can form a belief *partly* on the basis of experts' testimony and partly on the basis of other considerations. So, there is a question about how much of a role the expert testimony must play in a belief-forming process for it to be aptly characterized as a case of deference.

As a belief-forming process, deference to experts is supposed to contrast with trying to figure something out "for oneself"—with, in a recent slogan, "doing one's own research." And part of the reason for deferring *rather* than doing one's own research is that one recognizes one's own limitations in assessing complex data and evidence, and the experts' superiority in doing so.[3] Thus, merely taking expert testimony into account as one (perhaps minor) source of evidence among many is insufficient to count as deferring—even if one happens to arrive at the same view as the expert. One must arrive at this view because one afforded the expert testimony considerable weight.

Some philosophers have argued that one ought to defer to experts in a very strong sense of 'defer', whereby one *entirely* sets aside one's own judgments about the topic under consideration, and does not

---

[2] Notice that we are talking of deference to an expert's *testimony* rather than deference to their *judgments* or *beliefs*. Presumably in most cases we defer to expert testimony *because* we expect it to be a good proxy for the expert's judgments or beliefs, which we expect to be reliable. But forming a belief on the basis of expert testimony would still count as deference even if the testimony does not, as it turns out, reflect the expert's genuine beliefs. Indeed, there may even be some cases where we can expect an expert's testimony to be reliable even if they do not believe it (Lackey 2007). Hence, deference need not *always* be deference to an expert's judgment or beliefs. (Thanks to an anonymous referee here.) A trickier question is whether deference need always be deference to testimony. Perhaps there are some cases where we have access to an expert's judgment by means other than their testimony, and we can also be thought of deferring to them in these cases. If so, then deference could be understood as forming a belief on the basis of an expert's testimony *and/or* on the basis of their judgments (or other proxies for their judgments). For simplicity, we will just talk of deferring to expert testimony here, but readers can substitute this more complex characterization throughout if they prefer.
[3] For recent epistemological discussions of the "do your own research" slogan and its pitfalls, see Levy 2022; Ballantyne et al. 2024.

even try to gather or evaluate any (first-order[4]) data and evidence about it oneself—instead, quasi-automatically believing whatever the experts say (Huemer 2005; Zagzebski 2012; Constantin & Grundmann 2020; Grundmann 2021). This is often called the "preemption" view, because its proponents typically claim that expert testimony "screens off" or "preempts" any other evidence that laypeople might otherwise have, in effect neutralizing its normative force. Consequently, on this view, it would be a mistake to try to weigh the expert testimony against other reasons one has about the matter under consideration. But other philosophers have objected that the preemption view counsels believing the testimony of experts too uncritically. In its place they offer the "total evidence" view, according to which, though expert testimony should be afforded significant weight, it should still be assessed and weighed against the background of one's other evidence (Lackey 2018; Dormandy 2018; Jäger 2016).

We will not try to resolve the normative debate between the preemption and total evidence views here. Our interest in this section is in how to *define* the notion of 'deference'. Regardless of whether the preemption view is correct, we do not want to say that only the very strong kind of deference counselled by the preemption view *definitionally counts* as deference. Instead, we will require only that the expert's testimony is afforded "considerable weight." Our working definition, then, is this:

> **Deference to an expert (more precise definition).** A process of inquiry whereby one comes to believe the content of an expert's testimony because one affords the fact that the expert testified to it considerable weight.

The weight in question here is *evidential* weight. Though affording the expert's testimony considerable evidential weight needn't involve completely setting aside all of one's other reasons, it will involve taking the expert's testimony to be capable of outweighing many (apparent) competing considerations. Hence, it will also involve having *some* disposition to discount one's own judgments and not to get too deep into evaluating the (first-order) evidence for oneself.[5]

A crucial feature of deference to experts as we've defined it is that it is not merely an *outcome* where one's beliefs end up matching those that the experts espouse, but a *process of inquiry* (broadly construed[6]) that leads toward this outcome—one whereby one ends up with these beliefs (roughly) *because* the experts espouse them. Relatedly, deferring to an expert is distinct from listening to their arguments,

---

[4] As distinct from the evidence provided by the expert's testimony itself.

[5] This is of course still somewhat imprecise as to exactly how much weight one must afford the expert's testimony to count as deferring. This is perhaps tolerable, since deference plausibly comes in degrees—one can be more or less deferential to an expert depending on how much weight one assigns to an expert's judgment. However, one proposal that would make it somewhat more precise is that one must be treating the expert's testimony that *p* as giving one *sufficient* reason to believe *p*, at least in the case at hand. This is compatible with the possibility that under other conditions—such as ones where there are extremely strong countervailing reasons, or where there are strong independent reasons to distrust the expert—one would not take it as sufficient; hence, it is still not quite as strong as the kind of deferential attitude that the preemption view seems to counsel.

[6] Of course, it is not the sort of inquiry involved in doing one's research, where one gathers and evaluates first-order evidence for oneself.

evaluating these arguments for oneself, and being persuaded to agree.[7] Deference involves believing something *because an expert testifies to it*, and this does not require understanding or trying to evaluate the expert's reasons and arguments. All of this is important because it shows that existing empirical studies of mere *alignment* between the (espoused) views of laypeople and experts don't suffice to reveal whether laypeople *defer* to experts in our sense.

It is also important to distinguish deference to experts from the related but nevertheless distinct phenomenon of *trust* in experts. On most accounts, trust is an *attitude* that one takes toward the person or institution trusted (Jones 1996; Hawley 2014; Ngyuen 2022). By contrast, again, we are understanding deference as an inquiry-oriented *activity* or *process* of belief-formation. Thus, trust and deference are distinct in that trust causes or perhaps *disposes* one to defer, whereas deference is the *manifestation* of this disposition.[8] Consequently, one might in principle trust someone even if one never has occasion to defer to them.

   b) *What is an expert?*

Following Worsnip (forthcoming), we can distinguish "reliability" and "status" conceptions of expertise.[9] On the reliability conception, being an expert is a matter of being objectively more reliable (i.e. more likely to be right) than laypeople with respect to a particular domain or field of inquiry. On the status conception of expertise, by contrast, experts are those who have been accorded status as an authority on a topic—for example, through recognized credentials such as higher degrees, job titles, or being regularly consulted by the media on the topic in question.

We are not aware of any work systematically investigating how ordinary people deploy the concept of an 'expert', but anecdotally, both the reliability and status conceptions of expertise seem to be operative in everyday language on different occasions. Consider how someone skeptical of official authorities might either say "I don't trust the experts" or "those people aren't really experts," describing the very same people on both occasions. In the first usage, they employ the status conception—it can be paraphrased as "I don't trust the people *with the status of being experts*"—whereas in the second, they employ the reliability conception—it can be paraphrased as "those people aren't really *genuinely reliable* experts". Experts in the status sense might be referred to as 'so-called experts', whereas 'experts' in the reliability sense might be referred to as 'true experts' (hence, those who are thought *not* to be experts in the reliability sense might be said *not* to be 'true experts').

Both conceptions of expertise create problems for the simple advice that we should (always) defer to experts. The reliability conception makes it almost trivially true that deference to experts is a good

---

[7] Though there can be cases somewhat intermediate between the two, where one listens to the arguments to some extent but lends them more credence because they come from an expert. Thanks to an anonymous referee here.
[8] Unsurprisingly, empirical work suggests that when people hear conflicting claims from scientific experts, they agree more with the expert who seems more trustworthy to them (Gottschling et al., 2020).
[9] See Watson (2021, chapter 6) for another discussion of this distinction in terms of 'objective' and 'reputational' accounts of expertise. See also Grundmann (2018) for a finer-grained typology of expertise.

method of forming true beliefs. However, as we will explain in §2b, it can be very hard for laypeople to determine who is an expert in the reliability sense, which makes it hard for laypeople to follow the prescription to defer to experts in the reliability sense. On the other hand, it is much easier for laypeople to identify experts in the status sense (Anderson 2011), but it is much less obvious that deference to status experts is a good method of forming true beliefs, given that some research suggests that status can track numerous irrelevant factors such as physical attractiveness and social dominance (Kalick, 1988; Rahal et al., 2021).

It is important for work on deference to experts to specify whether it is employing the status or the reliability (or some other) conception, both so as to be clear about what follows from empirical findings and to avoid different authors talking past each other. For our purposes, we prefer to employ the status conception of expertise, as we think it opens up more productive lines of inquiry on our central empirical and normative questions. On the empirical side, it is easier for researchers to identify and classify status experts so as to empirically test whether people are willing to defer to them. On the normative side, precisely because it doesn't make deference to experts an obvious good in all situations, employing the status conception also allows us to open up interesting, non-obvious normative questions about *under what conditions* we are justified in refusing to defer to experts.

It is worth noting that on both conceptions, expertise is a domain-specific notion—one is an expert relative to a field of inquiry. However, often the boundaries of different subfields are not so clear-cut, and one may sometimes count as an expert (relative to laypeople) in fields that are adjacent to one's field of expertise. (For example, an oncologist is likely more knowledgeable than a layperson on simple questions in internal medicine.) Moreover, many questions of public interest seem to be *hybridized questions*, i.e., questions whose investigation requires combining evidence and methods across different fields (Ballantyne 2019a). In such cases, it may be hard to identify who the relevant experts are to whom one ought to defer. Moreover, with many complex policy issues, natural-scientific questions are intertwined with social-scientific ones. For example, pandemic policy debates (for example, about masking laws or vaccine mandates) turn both on natural-scientific questions in immunology and on social-scientific questions about, for example, which measures might be most likely to induce compliance among the general population. The difficulty in identifying a single clear domain of expertise for many issues of public relevance gives rise to the threat of *epistemic trespassing:* the worry that an expert in one field may cross into another domain where they lack the relevant expertise, and nonetheless make confident recommendations for laypeople to follow (Ballantyne 2019a).

2. Variables affecting the propriety and prevalence of deference to experts

Having clarified our subject-matter, we now turn to a survey of a number of variables that may affect either the (normative) *propriety* or the (empirical) *prevalence* of deference to experts. For each, we will attempt to integrate and compare work on the former and latter issue.

   a) *Laypeople's competence and capacities*

We begin with the competence[10] and capacities of the layperson who is (or is not) deferring. This includes (among other things) their prior knowledge relevant to a specific domain, their capacity to identify and gather (first-order) evidence in this domain, their competence in assessing that evidence and what it indicates, and their capacity to resist bias distorting their reasoning.

Intuitively, competence is clearly relevant to the propriety or rationality of deference. It stands to reason that other things equal, the less competent a layperson is with a topic, the more need they have to defer to experts on it.[11] However, it's worth noting that on the "preemption" view of rational deference (§1a), as long as the gap between the expert and the layperson is wide enough—as long as the expert is sufficiently more competent or reliable than the layperson—the precise degree of competence of the layperson is irrelevant to the degree to which they should defer to the expert. For, on this view, regardless of exactly how competent the layperson is, they should defer *entirely* to the expert, setting their own reasons completely aside. By contrast, on the total evidence view, the layperson can afford some weight to their own reasons and judgments, and plausibly, this weight should increase gradually as their competence and knowledge increases.

Interestingly, though, the limited social-scientific literature available suggests that a less intuitive relationship may hold: that is, the less educated laypeople are, the *less* likely they are report a tendency to defer to experts (Brossard & Nisbet 2007). This is the case even though laypeople, on average, do recognize that their own understanding is somewhat less than that of experts (Vaupotic et al., 2022).

What explains this? To begin with, people in general are prone to overestimate their own competence and underestimate their susceptibility to bias—and hence, their competence relative to that of experts.[12] This overestimation can come from many sources. For example, when scientific arguments are written in ways that are accessible to non-scientists, people are more likely to overestimate their knowledge, and less likely to defer to experts; this is known as the "comprehensibility effect" (Scharrer et al., 2014). Motivated reasoning—that is, the effect when one's reasoning is subconsciously biased by a desire; in this case, the desire to believe that one is competent—may also have a role to play in people's overestimation of their competence (Kunda & Sanitioso 1989; Dunning et. al. 1989). Moreover, some research supports the claim, known as the "Dunning-Kruger effect", that those who lack competence are more likely to overestimate their competence (Dunning 2011).[13] This may help to explain why those who lack competence are less likely to defer.

---

[10] In the psychological literature, 'competency' is often used to describe the state of people who are in pursuit of expertise—i.e., having competency is a state intermediate between being a novice and an expert (Alexander 2003). We are talking about laypeople who are not becoming experts and have no intention of doing so; thus we have used the term 'competence' to avoid confusion.

[11] At least, assuming that they cannot or will not simply develop the relevant competence. In other words, *holding fixed* a person's level of competence, less competence results in more need to defer. Thanks to an anomnymous referee for pushing us to clarify this.

[12] For a sample of the vast literature, see Ahlstrom-Vij 2013 and Ballantyne 2015.

[13] Though there has been recent controversy over whether this effect is real. See Gignac & Zajenkowski (2020) for concerns about it, and Dunkel et. al. (2023) for a qualified defense of it. See also Lackner et al. (2023) for a study suggesting that the relationship between knowledge and confidence is curvilinear, with people with intermediate levels of science knowledge being the most overconfident.

Thus, what matters to the likelihood of deference is not just the competence and knowledge a layperson has on a given subject, but also whether they have accurate *self-perceptions* of their competence and knowledge, and whether and how well their knowledge applies to a particular issue in a particular context (Barzilai & Chinn, 2018). Indeed, the ability to accurately assess one's own competence and knowledge (or lack thereof) is itself a kind of competence—or, in the nice phrase of Barzilai and Chinn (2018), an "epistemic meta-competence." It stands to reason that those who lack this meta-competence will be less likely to defer, or at least less likely to defer appropriately.

Moreover, appropriate deference to experts requires a number of other traits and abilities that are themselves best thought of as capacities or competences. First, it requires capacities to identify and control a variety of biases related to one's identity and self-concept, and to avoid and manage the threats to one's self-perception that can arise from deference. Ballantyne (2023) has argued that intellectual humility in general requires that people inhibit self-oriented motives in favor of an orientation toward reality and evidence. This type of inhibition is part of self-regulation, which is a set of knowledge, skills, and dispositions to activate, maintain, and refine goal-pursuit via active management of one's cognitive, motivational, emotional, and behavioral responses and actions (Greene et al., 2024). With respect to deference to experts in particular, Ballantyne has suggested that self-regulation is needed to inhibit temporary inclinations to "be knowledgeable" or "act like a well-educated person" in favor of a less egoistic but more reliable process for discerning truth.

Second, and relatedly, Koestner et al (1999) found that the way people conceive of autonomy affects their propensity to follow experts' advice when warranted. Some people have a "reactive" view of autonomy, on which it consists in being free of influence by others. On this conception, deference to experts is, by definition, not autonomous. Others have a reflective view of autonomy, on which autonomous decisions are guided by one's own values. This conception of autonomy makes more room for autonomous deference to experts, provided the decision to defer is itself made of one's own free will. Koestner and colleagues found that reflective autonomy predicted following experts' advice when warranted, whereas reactive autonomy predicted ignoring that advice. If we construe having a sophisticated, reflective view of autonomy as a capacity (in a broad sense), then it is another capacity that promotes deference to experts.

Third, appropriate deference requires at least some competence to determine which experts to defer to. These decisions require second-order evaluation strategies such as evaluating sources for reliability and trustworthiness (Bromme et al., 2015). As we will explore shortly in §2b, reliable experts may sometimes be hard to identify, and one may need some level of expertise oneself to effectively assess experts' track records. Again, this might help to explain why those with less competence—at least of a certain kind—will defer less often, and less appropriately.

In sum, there is a significant divergence between plausible normative ideals about how competencies and capacities *should* affect deference and the reality of how laypeople *do* defer (at least judging by our current empirical evidence). Thus, in addition to the need for more research regarding how laypeople's competencies and capacities affect their ability to make rational choices about whether and how much

to defer, additional scholarship is needed on how to help people refine and improve their competences and capacities for better deference decision-making.

b) *The competence of the expert*

Let us now turn from the competence of the layperson deferring, and toward the competence of the expert being deferred to, and how this affects the propriety and prevalence of deference. Again, this competence includes prior knowledge, capacity to identify and gather evidence, competence in assessing evidence, and capacity to resist bias. A good proxy for expert competence is susceptibility to error, and there is some psychological work on what explains experts' susceptibility to error (or lack of) in cases where this susceptibility to error is clearly measurable (e.g. Shanteau 1992; Kahneman & Klein 2009). In the context of deference, however, the key question is whether (and when) lay people are able to reliably assess the competence of experts. Just as it is highly intuitive that laypeople normatively ought to defer more the less competent they are, it is also highly intuitive that laypeople ought to defer more the *more* competent the expert is. However, this prescription is unhelpful without some way for laypeople to assess expert competence.

Research finds that laypeople do make quick and confident judgments of expert competence (Thon & Jucks 2017). However, one might wonder whether these judgments are, or even can be, well-founded. In a classic paper, Goldman (2001) lays out criteria for laypeople to look to in determining which experts are competent or, more broadly, reliable.[14] These criteria include the quality of experts' arguments, experts' past track records, experts' biases, and appraisals of experts by "meta-experts" who tell us which experts are most reliable (Goldman 2001: 93; see also Anderson 2011). Subsequent literature, however, has raised a number of problems about the extent to which laypeople can use these criteria to make informed judgments about expert competence (Millgram 2015: ch. 1 appendix A; Guerrero 2016; Ballantyne 2019b; Brennan 2020; Nguyen 2020; Levy 2022: 110-122). We will illustrate the problems by taking Goldman's criteria one by one.

First criterion: quality of experts' arguments

As we saw in §1a, the prescription to defer to experts gains much of its force from the fact that laypeople are often not sufficiently competent to evaluate the first-order data and evidence for themselves. But assessing the "quality of experts' arguments" seems very close to evaluating whether the first-order data and evidence really does support their views—precisely what those who need to defer to experts are *not* able to do effectively. As such, reasons to be suspicious of laypeople's capacity to assess that first-order evidence largely carry over to their assessments of experts' arguments. This

---

[14] Though this section will consider various criteria for assessing the reliability of experts, this does not amount to reverting to a reliability *conception* of expertise (§1b). We are still using the word 'expert' to refer to someone with a status of the expert. Nevertheless, one can ask, of any (status) expert, whether they are reliable. And the reliability of (status) experts might be one factor that affects the prevalence or propriety of deferring to them—if we can assess this reliability effectively. Someone employing the reliability *conception* of expertise could reframe what we say in this section not as raising questions about whether we can accurately assess experts' reliability, but rather about whether we can accurately determine *who is* an expert.

is intensified given that sound scientific reasoning is often nevertheless counterintuitive to the lay hearer (Young et al. 2018).

Goldman suggests that laypeople can at least assess how much experts have to say in favor of their views and how readily they seem to be able to rebut objections. But these surface phenomena are very imperfect guides to the actual quality of the experts' arguments. This is a good reason for experts to sometimes decline to "debate" non-experts—as, for example, when vaccine scientist Peter Hotez refused the invitation of Joe Rogan to debate Robert F. Kennedy, Jr. about vaccines. When laypeople have no way of assessing the evidential well-foundedness of competing claims, it is too easy for a talented debater who has the less well-supported side of the argument to come off as "winning" or at least forcing a "tie" simply by being able to generate *some* counterclaim in response to everything that is said, and by sounding confident and slick.

Second criterion: experts' track-records

The tendency to pay attention to past accuracy emerges early in development, as children face the developmental dilemma of relying on others to learn information about the world. For example, children are less likely to learn a new word from an adult who previously labelled an object incorrectly (e.g., labelling a ball a "shoe") compared to an adult with a record of accuracy (Koenig et al., 2004). This general sensitivity to past success persists in adulthood, as individuals attempt to distinguish between competent and incompetent experts based on their past reliability (Shanteau et al. 2002, Weiss & Shanteau 2003).

However, several problems may impede the ability of laypeople to form accurate impressions of an expert's track record. First and foremost, knowing whether an expert was right in the past requires not only knowing what they said in the past, but also some "independent" way of verifying whether what they said was true (i.e., one that doesn't itself take the expert's reliability for granted). For example, if an immunology expert now pronouncing on the efficacy of the COVID vaccine previously said that the H1N1 vaccine (released in 2009) was effective, then in order to count this as track record data in their favor, I must have some independent way (beyond the expert's own testimony, and that of others like them) of knowing that the H1N1 vaccine *was* effective. But often I don't have any independent way of verifying this. Again, for me to verify it in a way that was independent of expert testimony would often require me to assess the first-order evidence for myself, which is precisely what I'm not supposed to be doing when I'm required to defer.

There are some occasions upon which we do have independent ways of verifying experts' past performance and hence can at least in principle accumulate track record data. This is most obviously so when experts make straightforwardly verifiable predictions about future events, for example, when economists forecast whether inflation will increase over the next quarter (Tetlock 2017). Even in these cases, though, any kind of systematic track record data would require a lot of work to gather. Moreover, Huang (forthcoming) shows using a computational model that even with a God's eye view on expert track records, deferring to experts based on these track records can under certain "plausible" conditions impede the accuracy of a group's beliefs. This is both because the group may converge on favored experts *too quickly* before experts' track records are lengthy (making these track records a less

good guide to future performance), and because this reduces cognitive diversity (i.e. the variety of different thinking styles) within the group of experts being consulted and thus loses the (group) epistemic benefits that such cognitive diversity brings.

Third criterion: experts' biases

Laypeople often theorize about expert biases both on the left and right (for example, right-leaning laypeople often theorize that climate change scientists are biased in favor of left-wing conclusions, while left-leaning laypeople often theorize that economists are biased in favor of right-wing conclusions). However, again, it's not clear that laypeople are in a good position to responsibly make such judgments: they typically don't have insider access to the workings of scientific communities, let alone to the minds of individual scientists. There is ample empirical evidence of our unreliability in attributing biases to others, and indeed of our tendency to be biased ourselves in the way that we attribute bias (Ballantyne 2015; Stanovich 2021). As such, we should at least be cautious in trying to assess experts' reliability via our perceptions of whether they are biased.

Fourth criterion: meta-experts

Finally, the advice to use meta-experts to figure out which experts to trust immediately raises the question of which *meta*-experts to trust. This threatens a regress, since we'd then need meta-meta-experts to assess the meta-experts, and so on. Moreover, it's likely that those who are suspicious of the experts in a certain field also distrust the meta-experts who proclaim them reliable. For example, those who consider the *New York Times* an unreliable source of information probably also consider the fact-checking sites that rate the *New York Times* above *Fox News* to be unreliable (cf. Lynch 2020). So, there are also limitations to the usefulness of meta-experts in resolving debates about which experts to trust.

Summing up our review of all four of Goldman's criteria, it is questionable whether ordinary people are in a good position to make detailed judgments of putative experts' competence. If expert competence is typically not something we're able to assess very well, it may make less difference to the rationality of deferring to experts than one might first think.

What then? It may be that we have to be guided more by testifiers' *status* as experts than by detailed judgments of their reliability (cf. Anderson 2011). Laypeople do, in practice, rely on experts' formal credentials (educational degrees, formal positions, symbols of status, etc) as a proxy for competence. For example, one study found that people were more trusting of medical information when it came from those with medical credentials compared to those without (Thon & Jucks, 2017). There are limits to this, since well-credentialed individuals can nevertheless appear very untrustworthy to lay people (Willemsen et al., 2012). For example, people trust medical experts less when they use more jargon words (Thon & Jucks, 2017), and people sometimes trust those who have lots of "practical" experience over those with more theoretical academic credentials (Pew, 2020).

11

However, status may often be the best proxy we have for competence. Perhaps we should have a default of treating those with the status of experts as worthy of deference until we have positive reasons not to—that is, until there are factors that "defeat" the presumptive reason to defer that expert status brings. Moreover, some of the very factors that make it hard to assess expert competence also make it harder for this presumptive reason to defer to status-experts to be defeated. For example, the fact that we're so bad at making assessments of bias suggests that we should not be too quick to write off people with the status of experts on the basis that we think they're biased.

c) Ideological alignment

An oft-repeated claim in contemporary discourse about polarization is that people will disproportionately trust or defer to those who share their political outlooks. In the main, this claim is borne out by the empirical literature. Research suggests that people rate experts who share their prior views as more trustworthy than experts with differing views (Barzilai et al., 2020), are more likely to defer to experts who they perceive share their moral values (Johnson et al., 2018, 2021), and view experts as more credible when the experts frame issues in a way that resonates with their worldview (Lachapelle et al., 2014). A largescale analysis of Twitter users found that pro- and anti-vaccine individuals both trusted (purported) experts to an equal degree, but placed their trust in different experts (Harris et al., 2023). Similarly, belief in anthropogenic climate change is better predicted by partisan identity than by scientific knowledge (Kahan et al., 2012; Pew Research Center, 2016).

The power of ideology in driving deference was demonstrated in one study where participants were shown descriptions of Democratic or Republican welfare policies and were asked whether they agreed with them (Cohen, 2003). Crucially, the content of the policy was sometimes mismatched with the political label (e.g., a conservative policy labeled as "Democratic Policy"), but the participants tended to overlook the content of the policy and instead agree with whatever policy was labelled as being endorsed by their party. In general, liberals have more trust in institutions perceived as liberal-leaning, such as science, the media, and higher education, whereas conservatives have more trust in institutions perceived as conservative-leaning, such as law enforcement, businesses, and the military (Gallup, 2022).

Plausibly, these results are in part explained by the fact that people are inclined to trust or defer to those whom they perceive as having their best interests at heart, and people generally perceive their fellow ingroup members (where this includes those on their ideological "team") as being more benevolent (Tack et al., 2013). But it is also in part because they are, predictably, perceived as being more reliable (Park et al., 2016). And it may also be partly for reasons unrelated to any truth-oriented goal. Popular theories in psychology propose that individuals are motivated to adopt beliefs not merely for the purpose of discovering truth but also for social identity purposes—to achieve a sense of belonging and status in one's group (Van Bavel & Pereira, 2018).

Turning to more normative issues, it is often assumed in popular writing that the tendency to defer disproportionately to "co-partisans"—that is, to those who share one's ideology or values—is unprincipled or that it problematically contributes to polarization (Klein 2020, Stanovich 2021).

Partisan deference clearly occasions some social harms: for example, it may lead to the spread of misinformation within ideological communities as the truthfulness of a claim is superseded by other motivations to believe it (Van Bavel et al., 2024). However, some philosophers have argued that despite this, partisan deference can be rational from the individual perspective (Rini 2017, Levy 2022, Begby forthcoming).[15] Such philosophers typically appeal to one or more of three closely related points. The first is that we typically do in fact trust those who share our ideological outlook more than those who don't. The second is that at least *from my point of view*, those who share my ideological outlook are in fact more reliable about ideologically contested matters than those who are not. The third is that even if I shouldn't think that those who share my ideology are more likely than my political opponents to have *true beliefs*, it is reasonable for me to take my co-partisans to be less likely than my political opponents to deliberately try to *mislead* me. Together these three factors are claimed to make it rational to defer disproportionately to co-partisans.

Other philosophers, however, argue that it is irrational to defer on the basis of ideology (Worsnip 2019, 2022; Woodard forthcoming). First, they claim that merely trusting someone or *considering* them reliable does not suffice to make it rational to defer to them, because our states of trust and considering-reliable are themselves open to rational assessment. *Irrationally* trusting someone or *irrationally* considering them more reliable than someone else cannot make it rational to defer to them; this, to use terminology developed by philosophers in another context, would be problematic "bootstrapping". To bring this out, consider that if the reasoning provided by advocates of partisan deference were sound, then parallel reasoning would also suggest that if someone in fact trusts white people more than black people, or men more than women, they're rational to defer to the former more than the latter, thus licensing "testimonial injustice" (Fricker 2007).

Moreover, opponents of disproportionate deference to co-partisans offer some positive reasons to think that this practice is often irrational. As the literature on "irrelevant influences on belief" brings out (e.g. Cohen 2001; Vavova 2018), many of us know that our ideological commitments are shaped by arbitrary, contingent features of our upbringing and history. Had our upbringing and histories been different, we would plausibly have had different ideological commitments. This makes it hard to see why the fact that someone shares my ideological commitments, themselves a result of my contingent history, should be a marker of their reliability.[16] It's also questionable whether the fact that someone is on the same political "side" as me should mean that they are less likely to exploit or mislead me: politicians who seek their own electoral advantage can be incentivized to spread misinformation and

---

[15] Lepoutre (2021: ch. 6) argues for a more nuanced position: moderate partisanship in one's deference practices can be individually rational, while highly disproportionate deference to co-partisans is "epistemically irrational at an individual level [...but] fruitful at a broader, systemic level" (181) because of the way that dogmatism can stimulate discussion and the exchange of ideas. Our discussion here focuses on whether beliefs formed through disproportionate deference to co-partisans are themselves rational, not on whether this practice has long-run, collective advantages, even ones that might be termed "epistemic" in a broad sense. We also won't focus on non-epistemic advantages of deferring to co-partisans on individual occasions, such as the way it builds community or reduces the costliness of gathering information.

[16] There is some controversy about whether the realization that my belief is due to an irrelevant influence is in itself a reason to become less confident in it (see e.g., White 2010). However, the claim here is only that in light of the role that irrelevant influences play in determining our ideological commitments, it's dubious that the fact that someone shares my ideological commitments is a marker of their reliability. It's not clear that the considerations adduced against the former claim also challenge the latter.

whip up outrage among their own supporters. Whether this is as common as ideological opponents trying to mislead each other, however, is an empirical issue worthy of further study.

We have seen that there is a great deal of evidence that ideological alignment affects the tendency to defer. However, an important caveat is that there are at least some cases in which the perception that an expert or institution has *any* political alignment—even one aligned with the would-be deferrer's views—leads to distrust, compared with the perception that the expert or institution is politically neutral. Clark and colleagues (2023) found that liberals were less trusting of (and less willing to defer to) liberal-leaning institutions that seemed politicized (e.g., journalists) compared to those that seemed less politicized (e.g., physicists). Similarly, conservatives were less trusting of (and less willing to defer to) politicized conservative institutions (e.g., the Supreme Court) compared to less politicized institutions (e.g., firefighters). The effect of politicization also explains why people are generally more trusting of scientific experts than lobbyists, who seem to have an (even) greater political agenda (Konig & Jucks, 2019).

People distrust institutions the most when they are politicized and ideologically incongruent with their own views – when the scientific journal Nature endorsed Joe Biden in 2020, Trump supporters lost trust in the journal and in science in general (Zhang, 2023). This suggests a dilemma for scientists as to whether to support pro-science political campaigns, given the empirical fact that this engenders distrust in the objectivity of science by outgroup members.

d) Consensus and disagreement among experts

Another potentially relevant variable in determining the prevalence and propriety of deference is the extent to which experts agree (i.e., have consensus) or disagree about a given issue. Many philosophers have expressed the view (e.g. Anderson 2011, Huemer 2005, Guerrero 2016) that it is only when we know there is consensus about some issue that we ought to defer. When there is significant expert disagreement, by contrast, they think that we should suspend judgment. Part of the case for this is simply that when there is significant expert disagreement, there is no single "expert opinion" to defer to. In order to defer to experts in any straightforward sense, one would need to decide *which* disagreeing experts to defer to.[17] But as we've already seen in §§2b and 2c, there are reasons for skepticism about people's capacity to identify which experts to defer to, and about the practice of deciding which experts to defer to on the basis of ideological alignment. If this is so, there may be no way to responsibly defer in cases of expert disagreement.

Other philosophers (e.g. Goldman 2001; Brennan 2020) are more optimistic about our capacity to at least sometimes determine which experts are more likely to be reliable, and so deny the stringent view that deference can *only* be rational in cases of consensus. However, even if we take this more optimistic stance, disagreement among experts is plausibly at least some evidence of general expert unreliability.

---

[17] This is not to say that split expert opinion can't call for changes in a lay person's confidence in other ways. For example, one might start out very confident in something, discover that it's subject to expert disagreement, and then become less confident in it. This isn't *deference* to experts in the sense defined in this article, but it is clearly updating one's confidence in light of the total state of expert opinion. Thanks to an anonymous referee here.

The reasoning behind this contention starts with the point that when experts straightforwardly disagree about matters of fact, they cannot all be right. So if, for example, 50% of experts believe that inflation will go up in the next quarter and 50% of experts believe it will go down, at least 50% of the experts must be wrong, and so the probability of getting a correct answer from a randomly sampled expert within the group cannot exceed 0.5. If the same phenomenon plays out repeatedly among these experts on other questions, this suggests that the average reliability of members of the expert group cannot be very high. So widespread expert disagreement is at least some evidence of the unreliability of the experts on average.

Now, while it's *possible* that this unreliability is very unevenly distributed—one group of experts in the field is consistently reliable while another group is consistently unreliable—we should also account for the possibility that none of the experts in the field are consistently reliable. This provides reason to be reticent about deferring to *any* of these experts. By contrast, when the experts agree, we don't have the same evidence of unreliability, and hence don't have the same reason to be reticent about deferring. Thus, although there are explanations of expert consensus other than the hypothesis that they are all converging on the truth (Miller 2013), it's still true that all things equal, in conditions of expert consensus we have greater reason to think the experts are reliable and competent.

So much for whether people *ought* to defer more when there is consensus and less when there is disagreement; *do* they? Empirical work does seem to suggest that laypeople are more likely to defer when there is expert consensus. Bartos et al. (2022) found that informing respondents of the medical consensus concerning the safety of the COVID-19 vaccine was associated with increased vaccine uptake. Bialek et al. (2023) report similar findings: informing participants of the expert consensus regarding the safety of the COVID vaccine is associated with a modest improvement in attitudes towards the vaccine.[18] However, empirical work also suggests that laypeople do not always suspend judgment when experts disagree. Instead, they often attempt to "pick and choose" which experts to defer to—consonant with the findings we reviewed in §2c, often partly on the basis of which experts share their values (Johnson, Rodrigues & Tuckett 2021; cf. also Contessa 2022).

e) *Complexity of the topic*

In this subsection and the next, we turn to features of the topic being deferred about. One such feature worth attending to is a subject matter's *complexity*. While plausibly any subject matter that is appropriately considered to have experts is complex to some degree, it's also plausible that some topics are more complex than others. But just what is complexity? We propose the following tentative characterization. Topics are complex to the degree that they are comprised of complex *questions*. A question's being complex means that (i) it requires a large amount of evidence to justifiably answer –

---

[18] Johnston & Ballard (2016) present some complications, though. They find that people were more likely to trust claims associated with a consensus of the general public than they were to trust claims associated with a consensus of economists. It is possible that this effect is domain-specific, e.g. that the degree to which people trust doctors differs from the degree to which they trust economists.

call this the amount of evidence – and (ii) the relevant evidence is highly difficult to reliably evaluate – call this the difficulty of the evidence.

With a tentative characterization of complexity in hand, important empirical and normative questions concerning deference to experts arise. Empirically, are people more likely to defer to experts the more complex the subject matter? Normatively, is the rational pressure to defer to experts greater the more complex the subject matter? There has been little direct work on these questions. But we will try to draw some connections to existing literature that can perhaps help frame future work investigating the connections between complexity and deference.

Beginning with the empirical question, it has been found that when reading texts and then making a decision based on those texts, people's confidence in that decision was lower when the texts were complex (Scharrer et al., 2014). However, the complexity of a text is at best an indirect proxy for the complexity of the subject-matter under consideration, so this finding is only indirectly relevant to the topic at hand.

There is also a vast empirical and conceptual literature in psychology on types of knowledge (e.g., de Jong & Ferguson-Hessler, 1996; McCarthy & McNamara, 2021). Prominent typologies of knowledge include differentiating declarative knowledge (i.e., knowledge of simple facts), procedural knowledge (i.e., knowledge of how to do something), and conceptual knowledge (i.e., complex knowledge of how a phenomenon works or manifests). Different types of knowledge may tend to have objects with different degrees of complexity. Plausibly, as a very rough and defeasible generalization, conceptual knowledge tends to have more complex topics as its object. If so, and if people defer more the more complex the subject-matter, we might expect people to defer more about conceptual knowledge than simpler declarative or procedural knowledge. However, we have found no existing research on whether this is so.

Let's now turn to the normative question. As with the empirical question about the relationship between complexity and probability of deference, we have found no existing work directly about the relationship between complexity and *rational* pressure to defer. However, it's plausible that the more complex a domain of inquiry, the more rational pressure there is for a layperson to defer to experts. If the topic is more complex, then that means roughly that it requires more training, or experience, or background knowledge to understand. Training, experience, and background knowledge are precisely what laypeople tend to lack and experts tend to possess. For that reason, greater complexity of a topic may increase rational pressure to defer to the experts, as laypeople are less capable of working through the issues on their own.

So far, we have focused on the tendency and rationality of deference to experts on empirical issues, but we can raise these same questions regarding deference about normative issues.[19] To a first approximation, normative issues concern notions like what one ought or has reason to do, what is morally right or wrong, what would be good or bad,[20] and so on. This change in subject matter from the empirical to the normative brings with it a number of new considerations.

People do often seem to defer about normative (and, in particular, moral) matters: for example, children defer to adults about right and wrong (Piaget, 1932/2013); and many people similarly defer to religious authorities such as priests and rabbis. One factor that might affect whether people are likely to defer about moral matters in particular is whether they perceive morality as objective. The evidence about ordinary people's perceptions of the objectivity of morality is mixed. Some research suggests that morality is seen as being objective, notwithstanding the widespread disagreement about what the putatively objective moral facts are (Turiel et al., 1987; Shermer, 2016); while other research calls this into question (Sarkissian et al., 2011). One might speculate that the more objective people perceive morality to be, the more likely they are to defer about it as they might about other objective facts; though certain kinds of moral relativism—such as views on which moral truth is relative not to individuals but to whole cultures—might also be consistent with patterns of deference about moral questions.

As we saw in §2c, people are more likely to defer to experts who share their moral views. Importantly, this seems to apply even with regard to non-moral topics. On one possible model, we begin with initial alignment on normative issues and then begin deferring on non-moral topic. We might term this a "normative cascade." If someone thinks that Trump is correct in his morals, they might be more likely to defer to him about COVID treatment. Similarly, the fact that progressives tend to see science as a normative good and to trust the moral credentials of scientists might explain why they are more likely to defer to them and to think that they should take an active role in public policy debates (Pew Research Center, 2019).

Let's now turn to whether people *ought* to defer to experts about normative issues. This has been the subject of an extensive debate in the moral epistemology literature. This literature typically takes it for granted that we are required or at least permitted to defer to experts about empirical issues, and often claims an intuitive contrast in the moral case, where it is said that there is something strange or "off" about forming moral beliefs purely on the basis of testimony (Driver 2006, McGrath 2009, Hills 2009). For example, it's thought that it would be strange for me to believe that abortion is wrong (or, indeed, permissible) merely because an "expert" tells me it is, without their explaining any of the reasons why

---

[19] To be clear, we were previously addressing whether people *do* defer about *empirical* issues (under various circumstances) as well as whether they *ought* to defer about *empirical* issues. Now we are looking to the questions of whether people *do* defer about *normative* issues as well as whether they *ought* to defer about *normative* issues.

[20] Some philosophers use 'normative' more narrowly so that questions about goodness or badness are evaluative rather than normative. We will stipulatively use it in the broader sense here.

it is wrong (or permissible). Importantly, what is said to be strange is deferring about "pure" moral matters, i.e. moral question that do not themselves crucially depend upon empirical facts (McGrath 2009). It would not be strange for me to defer about an moral question that is "impure" in the sense that it does depend crucially on empirical facts, if what I'm really uncertain about (and seeking expertise on) is those empirical facts. For example, if a marine biologist and I share the same moral framework, I might defer to them on whether it is permissible to eat bivalves (such as oysters and mussels), if what I am really uncertain about (and seeking expertise on) is the empirical issue of whether bivalves feel pain. What would be (allegedly) odder would be if I knew all the empirical facts about whether bivalves feel pain or not, and then deferred to a distinctively moral expert about whether, given these facts, it is morally permissible to eat them.

Philosophers have given different explanations of what is strange about moral deference. Some think that it is because there simply *are* no (pure) moral experts, either in principle (because morality is inherently supposed to be "accessible" to everyone) or in practice (because those who might initially seem to be "moral experts", moral philosophers, disagree so extensively that they cannot be very reliable). Others argue that even though there might be moral experts, it is especially hard to identify who is a moral expert in a principled way (Nguyen 2020). Others offer other explanations of why it is problematic to defer about moral matters. For example, Hills (2009) argues that this is because moral beliefs formed on the basis of deference can't achieve moral *understanding*, and that moral understanding is required for acting with moral worth. Still others argue that in fact, moral deference *can* be permissible or even required after all, and so deny a significant discrepancy between the empirical and normative cases on this score (Sliwa 2012, Enoch 2014, Wiland 2017).

So far, we have focused on the tendency and rationality of deference when the topic is straightforwardly a normative one. We have thus relied on a distinction between deferring to experts about empirical matters and deferring to experts about normative matters. But, it is worth noting that the two may not be so cleanly separable. First, and most simply, experts about empirical matters do often give policy advice that turns (perhaps sometimes tacitly) not only on their empirical knowledge but also on normative judgments. Public health experts who study dangerous diseases also recommend that we should get vaccinated (CDC, 2023), and people who study inequality often chime in on economic policies; these recommendations reflect normative judgments about the proper goals of public policy and about the resolution of tradeoffs between different goods.

Moreover, there is an important debate in the philosophy of science concerning whether the enterprise of arriving even at apparently purely empirical scientific conclusions is itself necessarily value-laden. Those who argue that it is (e.g. Longino 1983, 1990; Douglas 2000, 2008, 2009) thereby challenge what is called the value-free ideal, according to which scientists' (non-epistemic) values should play no role in the justification of their conclusions. Douglas argues that science is value-laden on the grounds of what she calls inductive risk. Science, Douglas contends, requires drawing conclusions, and drawing a conclusion requires deciding what counts as sufficient evidence to do so. What counts as sufficient evidence to draw a conclusion depends at least partially on which result one takes to be worse in the

case at hand, a false positive or a false negative. And which is worse between a false positive and a false negative will depend at least partially on the scientists' values.

Douglas provides the following example. Suppose there is uncertain scientific evidence (as scientific evidence typically is) suggesting that the amount of some chemical deposited by a corporation into a nearby water supply can lead to an increased probability of cancer for those drinking the water. Whether the evidence counts as sufficient to draw the establish the relationship between the presence of the chemical and the increased probability of cancer depends at least in part on whether the scientists think (i) it is more important to reduce the risk of harming the health of the people drinking the water or (ii) it is more important to avoid overregulating business. (i) expresses a preference for false positives over false negatives. It's better to treat the chemical as being associated with an increased probability of cancer, and thus reduce the risk of harm to health, even if that ends up being incorrect and leading to overregulation. (ii) expresses a preference for false negatives over false positives. It's better to avoid overregulation, even if that leads to some harm to health.

Douglas' claims about the value-ladenness of science are not uncontroversial. Betz (2013) responds to Douglas' challenge to the value-free ideal by arguing that scientists can avoid importing value judgments if they make the uncertainties in the relevant evidence explicit when they present their findings, rather than making an all-or-nothing judgment on the basis of their values. Menon and Stegenga (2023) argue that even if value-free science is impossible to achieve, it is worth pursuing as an ideal. Moving beyond those who think that the value-free ideal is correct and attainable (or unattainable but a worthy ideal), there are those who accept Douglas' critique but assess more specifically what kinds of non-epistemic values can appropriately play a role in scientific inquiry. Intemann (2015), for instance, suggests that while non-epistemic values can appropriately play a role in scientific inquiry, those values should "promote democratically endorsed epistemological and social aims of research" (217).

Adjudicating the debate over the value-free ideal is beyond our scope. We bring it up to note that what comes out of that debate can potentially complicate any picture according to which laypeople ought to defer only with respect to empirical, but not normative matters. If scientific conclusions are often unavoidably influenced by scientists' values, then it may be that deference on empirical matters will involve deferring, in some sense, about the scientists' value judgments as well. We think this is an important line of future research concerning the rationality of expert deference.

3. The dark side of deference to experts

As we said at the outset, a popular narrative is that many people defer to experts too little. Our survey of the literature so far has considered the extent to which this is so by considering how different variables and how they affect the propriety and prevalence of deference to experts. On balance, the evidence does seem to suggest that people's practices fall short of the ideal in various ways. For example, we've seen that people are actually *less* likely to defer as they are less competent (when plausibly they ought to defer *more* when this is so); that people will decide not to defer to experts based

on negative evaluations of expert competence that they're arguably not in a good position to make; and that people are at least arguably unduly influenced by ideological alignment in their deference decisions. However, we also wish to sound some general notes of caution about the dangers of blanket deference to experts (and thereby of propounding simplistic norms that uncritically encourage this). We will explore four such dangers here.

a) Unwitting deference on value-questions

The first respect in which deference to experts might be dangerous follows on from the issues that we have just been discussing in §2f about moral deference and values. Even if there are moral (or, more broadly, normative) experts, they generally won't be scientists, even scientists with expertise on related empirical questions. For example, a climate scientist does not have any special expertise on the moral question of what our obligations to future generations are (cf. Broome 2012). An economic expert on what policies maximize GDP does not have special expertise on the moral/political question of whether GDP (as opposed to, say, income equality) is what ought to be maximized (cf. Barnes forthcoming). And an epidemiologist expert on how the COVID-19 virus spreads does not have special expertise on the question of how the risk of contracting the virus is to be balanced against the mental health risks of having very little social contact with others, or the risks of disruption to a child's education. While these scientists may of course have opinions about these normative issues, and legitimately so, there is no particular reason why we should *defer* to them about them. Moreover, when they present themselves as having *special* expertise about normative issues, they may be engaging in the aforementioned practice of "epistemic trespassing" (§1b).

Yet, particularly because (as we've noted) empirical and normative issues are so often difficult (arguably impossible) to disentangle, scientists' testimony, especially in the context of giving advice or making policy recommendations, will typically incorporate both empirical and normative elements. For example, consider an an epidemiologist who makes recommendations about what kind of lockdown policies should be implemented. In part, these recommendations are based on empirical judgments about what will contain the spread of COVID-19. But they are also (whether tacitly or explicitly) based on normative judgments. For while we may all agree that containing the spread of COVID-19 is important, there are difficult normative questions about how to weigh this goal against other important goals that it may trade off against (e.g., the mental health and educational costs of lockdowns). If the epidemiologist makes not just the judgment that a particular lockdown policy would help contain the spread of COVID-19 but the further judgment that this policy *overall* ought to be implemented, then they unavoidably presuppose some normative judgment about how these tradeoffs should be made.

Furthermore, it may not be transparent to everyone that these experts are making such normative judgments, still less what those normative judgments are. This means that when people defer to these experts about overall advice or recommendations, they may unwittingly defer to them not just about

empirical but about normative matters, despite the fact that such experts have no special expertise with respect to the latter.[21]

Some public narratives that encourage us to "follow the science" or "listen to the science" may encourage this unwitting deference to scientists about normative issues. Such narratives may misleadingly imply that science has normative policy consequences in a straightforward way, and fail to highlight the ways in which different values will yield different policy recommendations even holding the science fixed. Our public discourse about expertise often fails to distinguish empirical and normative questions, with people in effect being encouraged to defer to scientists about *both* kinds of issues (or, by others, about *neither*). If what we're suggesting is right, a better approach would be to defer to scientists about empirical but not about normative issues. Yet, again, because the two are often closely entangled in scientists' testimony—and are perhaps not even fully separable in principle—this is tough to do. This is a problem for deferring to scientific experts to the appropriate degree and in appropriate ways.

b) Confirmation bias

Confirmation bias refers to the tendency to more readily accept information that supports our pre-existing beliefs compared to information that contradicts it (Mynatt et al., 1977). This poses some problems for deference (and propounding norms encouraging it), especially because, as we saw in §3c, people tend to trust ideologically aligned experts. Our tendency toward confirmation bias raises the danger that when we try to follow the advice to "defer to the experts," we do so very selectively (Contessa 2022).[22] People may defer in an asymmetric way, increasing their confidence in beliefs when expert testimony confirms them, but failing to decrease confidence in their beliefs when expert testimony contradicts them.

Some of the most interesting work on the connection between morality and empirical facts concerns the "consequentialist crutch"—the idea that people will arrive at empirical beliefs that support their deontic convictions. For example, if you think that sex education (as currently practiced) is immoral, you will be motivated to conclude that it also does not reduce teen pregnancies (Ditto & Liu, 2012; cf. also Kahan & Braman, 2006). This suggests a sort of dual normative-empirical confirmation bias: we begin with strong moral intuitions about what is right, which leads us to seek out empirical evidence to confirm our moral convictions. Again, attempting to follow the advice to defer to experts can simply fuel these tendencies, when we seek out experts who provide evidence that confirms our existing beliefs.

c) Runaway echo chambers

---

[21] Arguably, public health experts might come closer to having both the requisite empirical *and* normative expertise. But as we saw in §2f, it's controversial whether there really are normative experts at all, and even if there are, it's an open question whether public health experts should be considered to be (reliable) experts about normative matters.
[22] Or, on a reliability conception of expertise (§1b), perhaps not at all.

Roughly speaking, an echo chamber is a structure in which a group's beliefs are continually reinforced while alternative perspectives and information are excluded or discredited (Nguyen, 2018). Echo chambers may not always lead us to false conclusions, if the sources we are exposed to within them are reliable. However, the nature of an echo chamber makes it hard to be rationally sure that these sources *are* reliable, precisely because competing perspectives are excluded or discredited (Sheeks 2023). Thus, even when one is lucky enough to form true beliefs in an echo chamber, these beliefs may not be rationally held (Sheeks 2023; Ranalli & Malcolm, 2023).

Once again, deference to experts within an echo chamber can be problematic. Echo chambers can produce incomplete or misleading perceptions of expert consensus, or lead us to defer to incompetent experts. This can lead to what Nguyen (2020) terms "runaway echo-chambers": when we start with false beliefs, we defer to incompetent experts who share those false beliefs, which reinforces our beliefs and also leads us to place greater faith in those (supposed) experts.

d) Anti-democratic attitudes

Finally, absolute, uncritical deference—of the sort arguably prescribed by the preemption view—has been thought by some to conflict with democratic values (Hazlett 2016; van Wietmarschen 2019). When our beliefs are effectively controlled by a relatively small group of experts, and our political behavior is influenced accordingly, we may seem to approach a kind of "epistocracy" or rule by experts (cf. Brennan 2016) rather than a society where citizens exercise their own autonomy and develop their democratic capacities in so doing. This might be particularly problematic when experts are demographically unrepresentative or don't share ordinary citizens' values. This problem interacts with the problem already surveyed in §3a—when scientists' judgments themselves encode value judgments and not merely empirical facts, giving them effective control over policy matters allows their value judgments, rather than those of ordinary citizens, to rule the day (cf. Harvard et al. 2021).

It is controversial both whether a society characterized by uncritical deference to experts really constitutes an epistocracy, and whether epistocracy is a bad thing. Interestingly, however, there is some empirical evidence that deferential attitudes to scientists are correlated with anti-democratic views about science policy—for example, the view that scientists should make policy decisions without consulting the public (Howell et al. 2020). While this is not fully dispositive, it is some evidence that the idea that there is a connection between (extreme forms of) deference and anti-democratic values is onto something.

4. Conclusion: directions for future research

Having surveyed existing normative and empirical work on a number of variables that might affect the prevalence and propriety of deferring to experts, we are now in a position to summarize some of the gaps in the existing literature that it would be good for future work to address.

First, as we pointed out in §1a, a great deal of the relevant existing work, especially empirical social scientific work, is not quite directly targeted at measuring willingness to *defer* in the sense that we have identified here, often instead mere *agreement* with experts.[23] This is true of much of the work that we summarized in §2, and in some cases we had to extrapolate likely implications for deference in particular. So, our first plea is for more empirical work that directly measures people's willingness to *defer* to experts—that is, to form the belief that the expert espouses *because* one affords the expert's testimony considerable weight.

Second, while there is a decent amount of normative and empirical work on the influence of the first three variables we surveyed in §2—laypeople's competence and capacities; expert competence; and ideological alignment; and, to a perhaps slighlty lesser extent, the fourth—consensus and disagreement among experts—we believe there is an especially strong need for significantly more work on the last two—complexity of the topic, and whether the issue is empirical or normative. As we noted in §2e, complexity is a variable that has received relatively little attention (in relation to deference to experts) on either the normative or the empirical side, and that warrants further investigation. And while there has been a lot of normative, philosophical work on potential asymmetries between deference about empirical and normative issues, we have found relatively little empirical work on whether people defer to experts less (or more!) on normative issues than on empirical ones. Given that we have stressed the dangers involved in unwittingly deferring to experts on normative matters about which they lack special expertise (§3a), we think that this work is particularly pressing. It could turn out that this is one dimension along which, contrary to the popular narrative, some people defer to experts *more* than is warranted.

Third, there is a variable that has received so little attention in the existing literature on deference that we did not even include it in our review, namely the layperson's level of *interest* in the topic concerned. This is importantly different from competence, since interest in a topic is no guarantee of competence regarding it. Our speculative hypothesis is that people tend to be less willing to defer about the issues that they are most interested in—that these are the issues about which they are motivated to "do their own research."

Yet we are skeptical that this is rational. As we noted in the Introduction, doing one's own research is in general a dangerous strategy when one lacks the competence to execute it effectively. Moreover, it is at its most dangerous when the stakes are high; that is, when the issue is one of great importance. Thus, to the extent that interest in an issue tracks its importance (at least to the interested party and those they care about), the normative case for deference (as opposed to doing one's own research) becomes *stronger* as interest increases, not weaker. We thus speculate that this may be a further respect in which the reality of people's deference practices diverges from the normative ideal. But further work on both the normative and empirical side of the issue is needed to see whether this speculation is correct.

---

[23] Work on *trust* in experts is closer to getting at deference, but as we pointed out in §1a, trust and deference are at least minimally distinct, and it is possible that they come apart in more radical ways.

Fourth, there are also interesting connections, worth investigating further, between the topic of deference to experts and deference to artificial intelligence. For instance, there has been much discussion about the importance of transparency and explainability of AI reasoning (Middlestadt et al. 2016, Zerilli et al. 2018). A central worry with artificial intelligence is that due to its nature as a 'black box,' we often cannot trace or understand its reasoning. And that raises concerns. If we don't understand how AI is reaching its conclusions, should we really be using those conclusions in our own reasoning and decision-making? Recently, there has been work comparing the problem of transparency as it pertains to AI to an analogous problem with expert reasoning. The worry is that we cannot track an AI's reasoning – but much the same can be said for the reasoning of experts in highly specialized fields (Ross 2022). Further exploration of what deference to experts and deference to AI have in common will hopefully help us better understand the rationality and of both.

Our review has explored when people in fact defer, when they *should* defer, and when the two come apart. One final and immensely important area for future research is to try to close the gap – to help us understand how we can successfully encourage deferring to experts when (and only when) it is rational to do so. Some of the work we have cited has gestured in this direction. Bartos et al. (2022), for instance, suggests that clearly communicating to people that there is expert *consensus* about an issue can promote deference. Clark et al. (2023) and Zhang (2023) suggest that taking explicit political stances can cause people to lose trust in experts. These studies thus offer some guidance on what *not* to do. But undoubtedly more research is needed to determine how we can further close the gap between when we should defer and when we do defer.

References

Ahlstrom-Vij, K. (2013). Why We Cannot Rely On Ourselves For Epistemic Improvement. *Philosophical Issues*, *23*(1), 276–296. https://doi.org/10.1111/phis.12014

Alexander, P. A. (2003). The Development of Expertise: The Journey From Acclimation to Proficiency. *Educational Researcher*, *32*(8), 10–14. https://doi.org/10.3102/0013189X032008010

Alexander, P. A. & The Disciplined Reading and Learning Laboratory. (2012). Reading Into the Future: Competence for the 21st Century. *Educational Psychologist*, *47*(4), 259–280. https://doi.org/10.1080/00461520.2012.722511

Anderson, E. (2011). Democracy, Public Policy, and Lay Assessments of Scientific Testimony. *Episteme*, *8*(2), 144–164. https://doi.org/10.3366/epi.2011.0013

Ballantyne, N. (2015). Debunking Biased Thinkers (Including Ourselves). *Journal of the American Philosophical Association*, *1*(1), 141–162. https://doi.org/10.1017/apa.2014.17

Ballantyne, N. (2019a). Epistemic Trespassing. *Mind*, *128*(510), 367–395. https://doi.org/10.1093/mind/fzx042

Ballantyne, N. (2019b). *Knowing our limits*. Oxford University Press.

Ballantyne, N. (2023). Recent work on intellectual humility: A philosopher's perspective. *The Journal of Positive Psychology*, *18*(2), 200–220. https://doi.org/10.1080/17439760.2021.1940252

Ballantyne, N., & Dunning, D. (2022). *Reason, Bias, and Inquiry: The Crossroads of Epistemology and Psychology*. Oxford University Press.

Ballantyne, N., Celniker, J. B., & Dunning, D. (2022). "Do Your Own Research." *Social Epistemology*, 1–16. https://doi.org/10.1080/02691728.2022.2146469

Barnes, G. (forthcoming). The Abuse of Expertise and the Problem with Public Economics. *Social Theory and Practice*.

Bartoš, V., Bauer, M., Cahlíková, J., & Chytilová, J. (2022). Communicating doctors' consensus persistently increases COVID-19 vaccinations. *Nature*, *606*(7914), 542–549. https://doi.org/10.1038/s41586-022-04805-y

Barzilai, S., & Chinn, C. A. (2018). On the Goals of Epistemic Education: Promoting Apt Epistemic Performance. *Journal of the Learning Sciences*, *27*(3), 353–389. https://doi.org/10.1080/10508406.2017.1392968

Barzilai, S., Thomm, E., & Shlomi-Elooz, T. (2020). Dealing with disagreement: The roles of topic familiarity and disagreement explanation in evaluation of conflicting expert claims and sources. *Learning and Instruction*, *69*, 101367. https://doi.org/10.1016/j.learninstruc.2020.101367

Begby, E. (2022). From Belief Polarization to Echo Chambers: A Rationalizing Account. *Episteme*, 1–21. https://doi.org/10.1017/epi.2022.14

Betz, G. (2013). In defence of the value free ideal. *European Journal for Philosophy of Science*, *3*(2), 207–220. https://doi.org/10.1007/s13194-012-0062-x

Białek, M., Meyers, E. A., Arriaga, P., Harateh, D., & Urbanek, A. (2023). COVID-19 vaccine skeptics are persuaded by pro-vaccine expert consensus messaging. *Journal of Experimental Psychology: Applied*, *29*(3), 477–488. https://doi.org/10.1037/xap0000467

Brennan, Jason. (2016). *Against democracy*. Princeton University Press.

Brennan, Johnny. (2020). Can Novices Trust Themselves to Choose Trustworthy Experts? Reasons for (Reserved) Optimism. *Social Epistemology*, *34*(3), 227–240. https://doi.org/10.1080/02691728.2019.1703056

Bromme, R., & Thomm, E. (2016). Knowing Who Knows: Laypersons' Capabilities to Judge Experts' Pertinence for Science Topics. *Cognitive Science*, *40*(1), 241–252. https://doi.org/10.1111/cogs.12252

Bromme, R., Thomm, E., & Wolf, V. (2015). From Understanding to Deference: Laypersons' and Medical Students' Views on Conflicts Within Medicine. *International Journal of Science Education, Part B*, *5*(1), 68–91. https://doi.org/10.1080/21548455.2013.849017

Brossard, D., & Nisbet, M. C. (2007). Deference to Scientific Authority Among a Low Information Public: Understanding U.S. Opinion on Agricultural Biotechnology. *International Journal of Public Opinion Research*, *19*(1), 24–52. https://doi.org/10.1093/ijpor/edl003

CDC. (2024). *Stay Up to Date with COVID-19 Vaccines*. Cdc.Gov. https://www.cdc.gov/coronavirus/2019-ncov/vaccines/stay-up-to-date.html

Clark, C., Isch, C., Everett, J., & Shariff, A. (2023). *Politicization Undermines Trust in Institutions, Even Among the Ideologically Aligned Public*. https://doi.org/10.21203/rs.3.rs-3239561/v1

Cohen, G. A. (2001). *If you're an egalitarian, how come you're so rich?* (4. print). Harvard Univ. Press.

Cohen, G. L. (2003). Party Over Policy: The Dominating Impact of Group Influence on Political Beliefs. *Journal of Personality and Social Psychology*, *85*(5), 808–822. https://doi.org/10.1037/0022-3514.85.5.808

Constantin, J., & Grundmann, T. (2020). Epistemic authority: Preemption through source sensitive defeat. *Synthese*, *197*(9), 4109–4130. https://doi.org/10.1007/s11229-018-01923-x

Contessa, G. (2022). Shopping for experts. *Synthese*, *200*(217). https://doi.org/10.1007/s11229-022-03590-5

De Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist*, *31*(2), 105–113. https://doi.org/10.1207/s15326985ep3102_2

Dormandy, K. (2018). Does Epistemic Humility Threaten Religious Beliefs? *Journal of Psychology and Theology*, *46*(4), 292–304. https://doi.org/10.1177/0091647118807186

Douglas, H. (2000). Inductive Risk and Values in Science. *Philosophy of Science*, *67*(4), 559–579. https://doi.org/10.1086/392855

Douglas, H. (2008). The Role of Values in Expert Reasoning. *Public Affairs Quarterly*, *22*(1), 1–18.

Douglas, H. E. (2009). *Science, policy, and the value-free ideal*. University of Pittsburgh Press.

Driver, J. (2006). Autonomy and the Asymmetry Problem for Moral Expertise. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *128*(3), 619–644.

Dunkel, C. S., Nedelec, J., & Van Der Linden, D. (2023). Reevaluating the Dunning-Kruger effect: A response to and replication of. *Intelligence*, *96*, 101717. https://doi.org/10.1016/j.intell.2022.101717

Dunning, D. (2011). The Dunning–Kruger Effect. In *Advances in Experimental Social Psychology* (Vol. 44, pp. 247–296). Elsevier. https://doi.org/10.1016/B978-0-12-385522-0.00005-6

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, *57*(6), 1082–1090. https://doi.org/10.1037/0022-3514.57.6.1082

Enoch, D. (2014). A Defense of Moral Deference: *Journal of Philosophy*, *111*(5), 229–258. https://doi.org/10.5840/jphil2014111520

Frankle, R. T. (1976). Nutrition education in the medical school curriculum: A proposal for action: a curriculum design. *The American Journal of Clinical Nutrition*, *29*(1), 105–109. https://doi.org/10.1093/ajcn/29.1.105

Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.

Gallup (2022). Confidence in US Institutions Down; Average at New Low. https://news.gallup.com/poll/394283/confidence-institutions-down-average-new-low.aspx

Gignac, G. E., & Zajenkowski, M. (2020). The Dunning-Kruger effect is (mostly) a statistical artefact: Valid approaches to testing the hypothesis with individual differences data. *Intelligence*, *80*, 101449. https://doi.org/10.1016/j.intell.2020.101449

Goldman, A. 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

Gottschling, S., Kammerer, Y., Thomm, E., & Gerjets, P. (2020). How laypersons consider differences in sources' trustworthiness and expertise in their regulation and resolution of scientific conflicts. *International Journal of Science Education, Part B*, *10*(4), 335–354. https://doi.org/10.1080/21548455.2020.1849856

Greene, J. A. (2022). What can educational psychology learn from, and contribute to, theory development scholarship?. *Educational Psychology Review*, *34*(4), 3011-3035. https://doi.org/10.1007/s10648-022-09682-5

Greene, J. A., Bernacki, M. L., & Hadwin, A. F. (2024). Self-Regulation. In P. A. Schutz & K. R. Muis (Eds.), *Handbook of educational psychology* (Fourth edition). Routledge, Taylor & Francis Group.

Grundmann, R. (2018). The Rightful Place of Expertise. *Social Epistemology*, *32*(6), 372–386. https://doi.org/10.1080/02691728.2018.1546347

Grundmann, T. (2021). Facing Epistemic Authorities: Where Democratic Ideals and Critical Thinking Mislead Cognition. In S. Bernecker, A. K. Flowerree, & T. Grundmann (Eds.), *The Epistemology of Fake News* (1st ed., pp. 134–155). Oxford University PressOxford. https://doi.org/10.1093/oso/9780198863977.003.0007

Guerrero, A. (2016). Living with ignorance in a world of experts. In R. Peels (Ed.), *Perspectives on Ignorance from Moral and Social Philosophy* (0 ed.). Routledge. https://doi.org/10.4324/9781315671246

Harris, M. J., Murtfeldt, R., Wang, S., Mordecai, E. A., & West, J. D. (2023). *The role and influence of perceived experts in an anti-vaccine misinformation community.* https://doi.org/10.1101/2023.07.12.23292568

Harvard, S., Winsberg, E., Symons, J., & Adibi, A. (2021). Value judgments in a COVID-19 vaccination model: A case study in the need for public involvement in health-oriented modelling. *Social Science & Medicine*, *286*, 114323. https://doi.org/10.1016/j.socscimed.2021.114323

Hawley, K. (2014). Trust, Distrust and Commitment. *Noûs*, *48*(1), 1–20. https://doi.org/10.1111/nous.12000

Hazlett, A. (2016). The Social Value of Non-Deferential Belief. *Australasian Journal of Philosophy*, *94*(1), 131–151. https://doi.org/10.1080/00048402.2015.1049625

Hills, A. (2009). Moral Testimony and Moral Epistemology. *Ethics*, *120*(1), 94–127. https://doi.org/10.1086/648610

Holiday, F. W. (1973). *The dragon and the disc: An investigation into the totally fantastic.* Norton.

Howell, E. L., Wirz, C. D., Scheufele, D. A., Brossard, D., & Xenos, M. A. (2020). Deference and decision-making in science and society: How deference to scientific authority goes beyond confidence in science and scientists to become authoritarianism. *Public Understanding of Science*, *29*(8), 800–818. https://doi.org/10.1177/0963662520962741

Huemer, M. (2005). Is Critical Thinking Epistemically Responsible? *Metaphilosophy*, *36*(4), 522–531.

Hutmacher, F., & Franz, D. J. (2024). Approaching psychology's current crises by exploring the vagueness of psychological concepts: Recommendations for advancing the discipline. *American Psychologist.* Advance online publication. https://doi.org/10.1037/amp0001300

Intemann, K. (2015). Distinguishing between legitimate and illegitimate values in climate modeling. *European Journal for Philosophy of Science*, *5*(2), 217–232. https://doi.org/10.1007/s13194-014-0105-6

Jäger, C. (2016). Epistemic Authority, Preemptive Reasons, and Understanding. *Episteme*, *13*(2), 167–185. https://doi.org/10.1017/epi.2015.38

Johnson, S. G. B., Rodrigues, M., & Tuckett, D. (2021). Moral tribalism and its discontents: How intuitive theories of ethics shape consumers' deference to experts. *Journal of Behavioral Decision Making*, *34*(1), 47–65. https://doi.org/10.1002/bdm.2187

Johnson, S., Rodrigues, M., & Tuckett, D. (2018). Cultural Values Guide Consumers' Deference to Experts. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3214001

Johnston, C. D., & Ballard, A. O. (2016). Economists and Public Opinion: Expert Consensus and Economic Policy Judgments. *The Journal of Politics*, *78*(2), 443–456. https://doi.org/10.1086/684629

Jones, K. (1996). Trust as an Affective Attitude. *Ethics*, *107*(1), 4–25. https://doi.org/10.1086/233694

Kahan, D. M., & Braman, D. (2006). Cultural Cognition and Public Policy. *Yale Law & Policy Review*, *24*, 147–170.

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, *2*(10), 732–735. https://doi.org/10.1038/nclimate1547

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, *64*(6), 515-526. https://psycnet.apa.org/doi/10.1037/a0016755

Kalick, S. M. (1988). Physical attractiveness as a status cue. *Journal of Experimental Social Psychology*, *24*(6), 469–489. https://doi.org/10.1016/0022-1031(88)90047-9

Klein, E. (2020). *Why we're polarized* (First Avid Reader Press hardcover edition). Avid Reader Press, an imprint of Simon & Schuster, Inc.

Koenig, M. A., Clément, F., & Harris, P. L. (2004). Trust in Testimony: Children's Use of True and False Statements. *Psychological Science*, *15*(10), 694–698. https://doi.org/10.1111/j.0956-7976.2004.00742.x

Koestner, R., Gingras, I., Abutaa, R., Losier, G. F., DiDio, L., & Gagné, M. (1999). To Follow Expert Advice When Making a Decision: An Examination of Reactive Versus Reflective Autonomy. *Journal of Personality*, *67*(5), 851–872. https://doi.org/10.1111/1467-6494.00075

König, L., & Jucks, R. (2019). When do information seekers trust scientific information? Insights from recipients' evaluations of online video lectures. *International Journal of Educational Technology in Higher Education*, *16*(1), 1. https://doi.org/10.1186/s41239-019-0132-7

Kröger, H., Donner, I., & Skiello, G. (1975). Influence of a new virostatic compound on the induction of enzymes in rat liver. *Arzneimittel-Forschung*, *25*(9), 1426–1429.

Kunda, Z., & Sanitioso, R. (1989). Motivated changes in the self-concept. *Journal of Experimental Social Psychology*, *25*(3), 272–285. https://doi.org/10.1016/0022-1031(89)90023-1

Lachapelle, E., Montpetit, É., & Gauvin, J. (2014). Public Perceptions of Expert Credibility on Policy Issues: The Role of Expert Framing and Political Worldviews. *Policy Studies Journal*, *42*(4), 674–697. https://doi.org/10.1111/psj.12073

Lackey, J (2007). Norms of Assertion. *Noûs*, *41*(4), 594-626. https://doi.org/10.1111/j.1468-0068.2007.00664.x

Lackey, J. (2018). Experts and Peer Disagreement. In M. A. Benton, J. Hawthorne, & D. Rabinowitz (Eds.), *Knowledge, Belief, and God: New Insights in Religious Epistemology* (pp. 228–245). Oxford University Press. https://doi.org/10.1093/oso/9780198798705.003.0012

Lepoutre, M. (2021). *Democratic Speech in Divided Times*. Oxford University Press.

Levy, N. (2022a). *Bad beliefs: Why they happen to good people*. Oxford University Press.

Levy, N. (2022b). Do your own research! *Synthese*, *200*(5), 356. https://doi.org/10.1007/s11229-022-03793-w

Liu, B., & Ditto, P. H. (2012). What Dilemma? Moral Evaluation Shapes Factual Belief. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2071478

Longino, H. (1983). Beyond "Bad Science": Skeptical Reflections on the Value-Freedom of Scientific Inquiry. *Science, Technology, & Human Values*, *8*(1), 7–17. https://doi.org/10.1177/016224398300800103

Longino, H. E. (1990). *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton University Press. https://doi.org/10.1515/9780691209753

Lynch, M. P. (2021). Polarisation and the problem of spreading arrogance. In A. Tanesini & M. P. Lynch (Eds.), *Polarisation, arrogance, and dogmatism: Philosophical perspectives*. Routledge.

McCarthy, K. S., & McNamara, D. S. (2021). The Multidimensional Knowledge in Text Comprehension framework. *Educational Psychologist*, *56*(3), 196–214. https://doi.org/10.1080/00461520.2021.1872379

McGrath, S. (2009). The Puzzle of Pure Moral Deference. *Philosophical Perspectives*, *23*, 321–344.

Menon, T., & Stegenga, J. (2023). Sisyphean science: Why value freedom is worth pursuing. *European Journal for Philosophy of Science*, *13*(4), 48. https://doi.org/10.1007/s13194-023-00552-7

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 2053951716679679. https://doi.org/10.1177/2053951716679679Miller, B. (2013). When is consensus knowledge based? Distinguishing shared knowledge from mere agreement. *Synthese*, *190*(7), 1293–1316.

Millgram, E. (2015). *The great endarkenment: Philosophy for an age of hyperspecialization*. Oxford University Press.

Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1977). Confirmation Bias in a Simulated Research Environment: An Experimental Study of Scientific Inference. *Quarterly Journal of Experimental Psychology*, *29*(1), 85–95. https://doi.org/10.1080/00335557743000053

Nguyen, C. T. (2020a). Cognitive islands and runaway echo chambers: Problems for epistemic dependence on experts. *Synthese*, *197*(7), 2803–2821. https://doi.org/10.1007/s11229-018-1692-0

Nguyen, C. T. (2020b). Echo Chambers and Epistemic Bubbles. *Episteme*, *17*(2), 141–161. https://doi.org/10.1017/epi.2018.32

Nguyen, C. T. (2022). Trust as an Unquestioning Attitude. In T. S. Gendler, J. Hawthorne, & J. Chung (Eds.), *Oxford Studies in Epistemology Volume 7* (1st ed., pp. 214–244). Oxford University PressOxford. https://doi.org/10.1093/oso/9780192868978.003.0007

Pew Research Center (2015). The Politics of Climate. https://www.pewresearch.org/science/2016/10/04/the-politics-of-climate/

Pew Research Center (2019) Trust and Mistrust in Americans' Views of Scientific Experts. https://www.pewresearch.org/science/2019/08/02/trust-and-mistrust-in-americans-views-of-scientific-experts/

Pew Research Center (2020). Science and Scientists Held in High Esteem Across Global Publics. https://www.pewresearch.org/science/2020/09/29/science-and-scientists-held-in-high-esteem-across-global-publics/

Piaget, J. (2013). *The moral judgment of the child.* Routledge. (Originally published 1932).

Rahal, D., Fales, M. R., Haselton, M. G., Slavich, G. M., & Robles, T. F. (2021). Cues of Social Status: Associations Between Attractiveness, Dominance, and Status. *Evolutionary Psychology*, *19*(4), 14747049211056160. https://doi.org/10.1177/14747049211056160

Ranalli, C., & Malcom, F. (2023). What's so bad about echo chambers? *Inquiry*, 1–43. https://doi.org/10.1080/0020174X.2023.2174590

Rini, R. (2017). Fake News and Partisan Epistemology. *Kennedy Institute of Ethics Journal*, *27*(2S), E-43-E-64. https://doi.org/10.1353/ken.2017.0025

Ross, A. (2024). AI and the expert; a blueprint for the ethical use of opaque AI. *AI & Society 39*(3), 925–936. https://doi.org/10.1007/s00146-022-01564-2

Sarkissian, H., Park, J., Tien, D., Wright, J. C., & Knobe, J. (2011). Folk Moral Relativism. *Mind & Language*, *26*(4), 482–505. https://doi.org/10.1111/j.1468-0017.2011.01428.x

Scharrer, L., Stadtler, M., & Bromme, R. (2014). You'd Better Ask an Expert: Mitigating the Comprehensibility Effect on Laypeople's Decisions About Science-Based Knowledge Claims: Mitigating the comprehensibility effect. *Applied Cognitive Psychology*, *28*(4), 465–471. https://doi.org/10.1002/acp.3018

Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational Behavior and Human Decision Processes*, *53*(2), 252-266. https://doi.org/10.1016/0749-5978(92)90064-E

Shanteau, J., Weiss, D. J., Thomas, R. P., & Pounds, J. C. (2002). Performance-based assessment of expertise: How to decide if someone is an expert or not. *European Journal of Operational Research*, *136*(2), 253–263.

Sheeks, M. (2023). The Myth of the Good Epistemic Bubble. *Episteme*, *20*(3), 685–700. https://doi.org/10.1017/epi.2022.52

Shermer, M. (2016). Morality is real, objective, and natural. *Annals of the New York Academy of Sciences*, *1384*(1), 57–62. https://doi.org/10.1111/nyas.13077

Sliwa, P. (2012). In defense of moral testimony. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *158*(2), 175–195.

Stanovich, K. E. (2021). *The Bias That Divides Us: The Science and Politics of Myside Thinking*. The MIT Press. https://doi.org/10.7551/mitpress/13367.001.0001

Tetlock, P. E. (2017). *Expert political judgment: How good is it? How can we know?* (New edition). Princeton University Press.

Theodor, J. L., & Senelar, R. (1975). Cytotoxic interaction between gorgonian explants: Mode of action. *Cellular Immunology*, *19*(2), 194–200. https://doi.org/10.1016/0008-8749(75)90203-8

Thon, F. M., & Jucks, R. (2017). Believing in Expertise: How Authors' Credentials and Language Use Influence the Credibility of Online Health Information. *Health Communication*, *32*(7), 828–836. https://doi.org/10.1080/10410236.2016.1172296

Turiel, E., Killen, M., & Helwig, C. C. (1987). Morality: Its structure, functions, and vagaries. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children* (pp. 155–243). University of Chicago Press.

Van Bavel, J. J., & Pereira, A. (2018). The Partisan Brain: An Identity-Based Model of Political Belief. *Trends in Cognitive Sciences*, *22*(3), 213–224. https://doi.org/10.1016/j.tics.2018.01.004

Van Wietmarschen, H. (2019). Political testimony. *Politics, Philosophy & Economics*, *18*(1), 23–45. https://doi.org/10.1177/1470594X18798062

Vaupotič, N., Kienhues, D., & Jucks, R. (2022). Gaining insight through explaining? How generating explanations affects individuals' perceptions of their own and of experts' knowledge. *International Journal of Science Education, Part B*, *12*(1), 42–59. https://doi.org/10.1080/21548455.2021.2018627

Vavova, K. (2018). Irrelevant Influences. *Philosophy and Phenomenological Research*, *96*(1), 134–152. https://doi.org/10.1111/phpr.12297

Watson, J. C. (2021). *Expertise: A Philosophical Introduction*. Bloomsbury Academic. https://doi.org/10.5040/9781350083875

Weiss, D. J., & Shanteau, J. (2003). Empirical Assessment of Expertise. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*(1), 104–116. https://doi.org/10.1518/hfes.45.1.104.27233

Whitcomb, D., Battaly, H., Baehr, J., & Howard-Snyder, D. (2017). Intellectual Humility: Owning Our Limitations. *Philosophy and Phenomenological Research*, *94*(3), 509–539. https://doi.org/10.1111/phpr.12228

White, R (2010). You Just Believe That Because… *Philosophical Perspectives*, *24*(1), 573-615. https://doi.org/10.1111/j.1520-8583.2010.00204.x

Wiland, E. (2017). Moral Testimony: Going on the Offensive. In R. Shafer-Landau & E. Wiland (Eds.), *Oxford Studies in Metaethics* (Vol. 12). Oxford University Press. https://doi.org/10.1093/oso/9780198805076.003.0003

Willemsen, L. M., Neijens, P. C., & Bronner, F. (2012). The Ironic Effect of Source Identification on the Perceived Credibility of Online Product Reviewers: The Ironic Effect of Source Identification. *Journal of Computer-Mediated Communication*, *18*(1), 16–31. https://doi.org/10.1111/j.1083-6101.2012.01598.x

Woodard, E. (forthcoming). What's Wrong with Partisan Deference? *Oxford Studies in Epistemology*, 8.

World Health Organization. (2020). *Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation*. https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation

Worsnip, A. (2019). The Obligation to Diversify One's Sources. In C. Fox & J. Saunders (Eds.), *Media ethics, free speech, and the requirements of democracy*. Routledge.

Worsnip, A. (2022). Review of *Bad Beliefs: Why They Happen to Good People* by N. Levy. *Notre Dame Philosophical Reviews*.

Worsnip, A. (forthcoming). Deference to Experts. In J. Darcy, E. Sosa, M. Steup, & K. Sylvan (Eds.), *A Companion to Epistemology* (3rd ed.). Wiley-Blackwell.

Wrightstone, R. N., Smith, L. L., Wilson, J. B., Vella, F., & Huisman, T. H. (1975). Some physicochemical properties of hemoglobin-manitoba (alpha2 102Ser replaced by Arg (G9) beta2). *Biochimica Et Biophysica Acta*, *412*(2), 283–287. https://doi.org/10.1016/0005-2795(75)90042-2

Young, A. G., Laca, J., Dieffenbach, G., Hossain, E., Mann, D., & Shtulman, A. (2018). Can Science Beat Out Intuition? Increasing the Accessibility of Counterintuitive Scientific Ideas. *CogSci*. Proceedings of the Annual Meeting of the Cognitive Science Society. https://sites.oxy.edu/shtulman/documents/2018d.pdf

Zagzebski, L. T. (2012). *Epistemic authority: A theory of trust, authority, and autonomy in belief*. Oxford University Press.

Zerilli, J., Knott, A., Maclaurin, J., & Gavaghan, C. (2019). Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard? *Philosophy & Technology*, *32*(4), 661–683. https://doi.org/10.1007/s13347-018-0330-6

Zhang, F. J. (2023). Political endorsement by Nature and trust in scientific expertise during COVID-19. *Nature Human Behaviour*, *7*(5), 696–706. https://doi.org/10.1038/s41562-023-01537-5