**Expertise and the Ethics of Trust:A Review**

**Maria Baghramian and Silvia Caprioglio Panizza**

**Abstract**

This review paper on the ethical issues surrounding trust in experts focuses on epistemic trust and its ethical implications. The paper is divided into four sections: Section I maps key philosophical discussions of trust; Section II discusses the breakdown of trust; Section III focuses on the core issue in the ethics of trust and is divided into four subsections; Section IV deals with the question of distrust. We conclude with some suggestions for creating conditions in which both trust and distrust can play socially, epistemically, and morally useful roles within a broader climate of trust.

***Keywords*: Ethics, Trust, Expertise, Distrust, Vulnerability.**

**Introduction**

Experts and expertise, their roles and scope, have come under a significant academic and political scrutiny. In our daily lives, we routinely depend on various , their skills, and their advice. From matters of health to technology, weather forecasts to air travel, and even in mundane matters of dealing with blocked drainpipes or broken washing machines, experts play a ubiquitous role in our lives and guide our choices. Experts and our reliance on them gain political significance when they begin to occupy a significant role in policy formation and implementation.

Recent years have seen a sharp turn toward populist politics, with leaders claiming to represent the univocal "will of the people" and to stand against the "liberal elites" and the privileged. Cosmopolitan. educated classes (e.g. Canovan, 1999). This anti-elitist rhetoric has put the scientific advisory process under serious stress (OECD, 2015). Policy advice on health and environmental issues has proven particularly controversial and led to partisan political debates and confrontations across the world. While headline figures in surveys on trust in scientific expertise do not indicate a drop in trust levels—unlike trust in politicians and the media—the public discourse around expertise has noticeably changed, and there is evidence of a breakdown of trust in specific policy areas such as vaccination and climate change (Kelsall & Sorell, 2024; Facciolà et al., 2019).

The role of experts in policy decisions and the question of trust that the general public is to place on them have far reaching ethical consequences. The current paper is an attempt to highlight some key points of the requirements of trust in experts and their ethical implications.[[1]](#footnote-1)

**Section I: Mapping Trust**

Epistemic dependence on experts is bound up with the question of trust. The peculiar form of trust at stake in discussing the relationship between experts and non-experts is epistemic trust or trust that applies to agents’ beliefs and the reasons provided for their beliefs, rather than their actions (Hardwig, 1991, p. 697). Epistemic trust in general, and trust in experts in particular, often take the form of testimonial trust, i.e. trusting what the experts tell non-experts, as well as the policies that are based on these recommendations. In this section, we inquire into the nature of the trust-attitude that non-experts should have toward experts.

We begin by considering some definitions of epistemic trust. Section II provides a survey of the ethical features of trust; Section III investigates the role of distrust in the understanding of the ethical requirements of trust. Three main families of views—predictive accounts, normative accounts, and combined accounts (Dormandy, 2020)—dominate current discussions of the philosophy of trust. Predictive accounts argue that to trust someone amounts to forming a positive expectation—if not a belief—that they will behave as agreed or required by the situation (Hardin, 1993). Normative accounts, by contrast, argue that to trust someone involves not just a prediction, but a normative demand about how the trustee ought to behave (Holton, 1994; Jones, 1996; Faulkner, 2007; Darwall, 2017). A weak version of this view concedes that one can trust another while suspending judgment about the likelihood that they will behave as expected, but not that one can have negative expectations toward the trustee (e.g., Holton, 1994). In contrast, the strong version of the normative account is compatible with negative predictive expectations, in that all it takes for one to trust another is to place normative expectations on their behavior (Jones, 2004).

We will return to the normative accounts of trust in our discussion of the ethics of trust in Section II, but on the face of it, when it comes to trust in experts in the policy domain, it looks as though a purely normative account is a non-starter, in that such accounts concede that one can trust another even when one believes that the trustee will not act as agreed or expected. We can make sense of this account in the context of a relationship, say, between parents and children, but it cannot apply to the domain of policy expertise, where no institution would request consultation from someone they consider unable or unwilling to deliver the requested outputs. To put it differently, there seems to be little room for *therapeutic trust* in the context of expert advice and policymaking (Faulkner, 2007; Nickel, 2007).

The predictive account of trust fares better because it accommodates the intuitive idea that non-experts and institutions select experts to provide policy advice based on the experts’ credentials and a considered—and likely reliable—expectation that such experts are able and willing to fulfil their function. Crucially, they predict that the experts will deliver the expected results. Predictive expectations take numerous forms, ranging from placing a high degree of confidence in the information provided, relying on the information provided, ascribing credibility to sources of information (including the person testifying) and having justified expectation of accuracy, usually cashed out in terms of truth. What these kinds of predictive expectations have in common is their epistemic goal—namely the production of epistemic goods such as knowledge, justified beliefs, understanding, and inquiry (Grasswick, 2020).

One reason we might want to go beyond a purely predictive view of trust, even in the policy domain, is that this view reduces trust to mere reliance (e.g., Goldberg, 2020), where the latter—unlike the former—requires no commitment on the part of the trustee to display an appropriate reaction to the trustor’s attitude. Yet, it could be argued that when non-experts and institutions put their trust in a policy advisor, they expect both that the trustee will act as predicted and that the trustee will do so because of a normative stance that the trust relationship creates.

This normative stance can be cashed out in various ways. Some regard it as an expectation of the trustee’s goodwill (Almassi, 2012; Baier, 1986; Cogley, 2012; Frost-Arnold, 2013; Wilholt, 2013); others, as a participant stance according to which the trustor treats the trustee as a person who bears responsibility for their actions (Holton, 1994); and still others, as a mere responsiveness of the trustee to the fact that the trustor is counting on them to act as expected (Faulkner, 2017). Normative and affective expectations, unlike predictive ones, lead to feelings of betrayal and not just disappointment when trust is broken (Baier, 1986, p.285). A plausible way to account for this normative dimension of a trust relationship in the context of expert advice involves requiring that experts at least comply with the ethical and epistemic norms of scientific practice, or that they display honesty, integrity, and other moral-epistemic virtues expected of scientists and researchers. A willingness to act in the interest of the recipients of their advice is also seen as a feature of the integrity expected of experts, and such willingness is taken as an indicator of their benevolence (see Hawley, 2017, for contrary view). These normative and affective expectations are the reassurances that we need in the face of the risks we take in trusting and justify the hope and confidence we place on those we trust. As we will see in the next section, not fulfilling such expectations is one of the reasons for the breakdown of trust in experts.

Trust is an essential element of societies, and especially of democracies: it creates cohesion and the very possibility of successful communication—particularly when the goal is for the public to take some action (as in the case of vaccination campaigns or changes in consumption patterns in response to climate change), or to accept certain changes that will impact their lives (for instance, again in the context of climate change, accepting increased taxes on highly polluting products or policy requirements to transition business production methods to low-carbon alternatives ). Currently, two particularly relevant domains are the communication of expert knowledge in the medical and environmental fields. In Section IV, in the course of discussing the ethics of trust, we shall consider the ethical implications of trust and distrust in the aforementioned cases.

**section II. The Breakdown of Trust**

The breakdown of trust in experts is a source of socio-political concern and a topic of philosophical interest. This section briefly examines some of the reasons for withdrawals of trust from experts. Mistrust of experts, like trust, has many sources and explanations. Trust can justifiably be withdrawn from experts who are judged to have made serious mistakes or have been dishonest, untruthful, or biased. The legitimacy of these concerns, at least in principle, is acknowledged by experts and non-experts alike. But it is easy to imagine that adequate training, professional vigilance, and public monitoring of the markers of intellectual *integrity—*as well as vigilance regarding the institutional norms governing the work of expert bodies(e.g., rigorous review mechanisms, political independence, etc.)—*could address* such concerns. Serious disagreement among experts is thought to have an impact on the perceptions of the trustworthiness of their advice, but the exact scope of this concern is in question (Dellsén, 2018).

The question remains why large numbers of people reject scientific consensus on crucial issues like climate change (in the United States) and vaccination (in the United Kingdom and the United States) Scepticism about expert advice in such cases rarely comes down to the details of the scientific evidence or the methodology scientists employ, but is linked to social, psychological, and broadly normative considerations (Levy, 2019).

Let us consider the psychological aspects first. At an individual level, traditional cognitive biases such as confirmation bias, desirability bias, and motivated reasoning obstruct trust relationships whenever there is a clash of opinions between the expert and the lay person (e.g., Nichols, 2017). At a collective level, the opportunity to establish immediate connections with people who share one’s worldview—typically online, via social media—makes novices prone to group polarization, that is, the tendency to adopt more extreme beliefs when engaging with a group of like-minded individuals (Sunstein, 2017). Combining these factors, we can easily make sense of Kahan and colleagues’ *cultural cognition thesis*, namely, that people tend to form beliefs about societal risks and factual information that sustain their personal values, and of *politically motivated reasoning*, that is, the idea that people trust experts who appear to share their values and distrust those who seem to hold diverging views (Kahan et al., 2010). The icing on the cake is offered by the *Dunning-Kruger* *effect* (Kruger and Dunning, 1999), that is, the tendency of novices to overestimate their ability in a given domain. As Ballantyne (2019) and Brennan (2020) point out, this psychological phenomenon is particularly relevant in cases where laypeople seem unable to acknowledge who is epistemically superior in a given domain.

Socio-epistemic structures such as epistemic bubbles and echo chambers, which reinforce ideological exclusion in different ways (Nguyen, 2018), strengthen psychological support for mistrust. Their combination leads us to increase in in-group trust and higher levels of distrust of outsiders, regardless of their expertise.

A further, but no less threatening aspect, has to do with the suspicion, if not outright disdain, shown toward experts and their advice by populist politicians such as Donald Trump in the United States or Bolsonaro in Brazil. The negative attitude of populist leaders toward experts is unsurprising. Populists seek to govern directly, establishing an unmediated, emotionally replete bond with the “real people.” Experts, with their evidence-based policy recommendations, aspirations of cool-headed objectivity, high educational achievements, and unabashed desire to be among the elite in their field, stand in stark contrast to the populist vision of politics and thus become ready targets of their ire. It should be noted that even populist regimes have to rely on experts, but their choice of experts frequently reflects their political stance and frequently elevates fringe scientists.

There are, however, a number of broader ethical concerns around the role of experts in our personal and public life. We will return to this in the next section.

**Section III. The Ethics of Trust**

Trust is often used as an ethically loaded concept, and trustworthiness is typically understood as a morally desirable property. It is not surprising, therefore, that trust is an important and complex concept in ethics. Trust is also considered to be not only morally relevant, but an *essential* element of morality. Thus Annette Baier, one of the pioneers of the ethics of trust, claims that “some degree of trust in the social world is the starting point and very basis of morality” (Baier, 2004, 180). The ethical dimension of trust depends on three key factors: the vulnerability that trust creates in the trustor and the corresponding possibility of harm for the trustee a result of trust; the related moral obligations that trust generates— either a responsibility to trust in certain situations, or obligations created in the trustee by contexts in which they are expected to prove trustworthy, given their power over the trustor; and the virtue of trustworthiness, along with the related moral virtues required by trust.

And yet, trust is not always—or uncontroversially—considered morally good.While trust and trustworthiness are more often intuitively associated with positive ethical value, there are also moral perils that accompany trust. In the final part of Section II, we turn to problematizing trust and asking whether there is any indication in the literature that trust sometimes ought to be withheld—even when one simultaneously has good reasons to trust— or that trust may create morally undesirable situations.

**III.1 The Value of Trust: Intrinsic and Instrumental Value**

In order to situate the discussion and highlight the overall moral significance of trust, and before investigating the specific moral dimensions of trust and trust relationships—the virtues, obligations, and responsibility of trustors and trustees—we shall begin by exploring the value of trust itself. It should be emphasized that the arguments presented here for the value of trust refer to justified trust (with the possible exception of therapeutic trust). While most of the literature focuses on the instrumental value of trust—the many ways in which trust can bring about morally desirable outcomes—some have also argued for the intrinsic value of trust. Colin O’Neil (2012), for example, aims to provide an account of the significance of trust in contrast with lying, betrayal, and the abuse of trust. When individuals succeed in exploiting trust in communication, they abuse trust. According to O’Neil, such abuses of trust violate the *gratitude* that trust demands, and it is in this gratitude, he argues, that the distinctive moral value of trust lies.

More has been said about the instrumental value of trust, ranging from trust’s ability to contribute to ethical and social goods, to the stronger idea that trust is fundamental to some of them. Cooperation, for instance, is impossible without trust, according to Friedrich and Southwood (2011), because of the necessity of trust in the promising involved in cooperation. The idea that trust is necessary to cooperation leads directly to the value of trust for social cohesion (to be discussed further below), seen as one of the central goods for which trust is instrumentally valuable. Hence, Fukuyama (1996) has observed how societies with a higher degree of trust are stronger, while Skyrms (2008) stresses the fact that mutual trust enables the “social contract” which, according to the theory, is the foundation of organized society. Conversely, Alfano (2016), in his investigation of “communities of trust,” shows how the absence of trust is linked to the absence of important social and individual goods.

Cooperation typically requires mutual trust, and that is what makes trust so valuable to a society. These dynamics can be observed at all levels and in different groups. More broadly, it can be said that trust is a key element of all relationships, and that the degree of trust that is present will influence the strength and stability of that relationship. Thus, Matthew Harding (2011), in comparing trust with respect, argues that trust’s demands can be unlimited, while respect makes limited demands—and this, he suggests, is one of the reasons why trust can go further than respect in building relationships.

The value of trust is not limited to mutuality. More difficult but equally important are cases in which the majority of the burden lies on one party, such as relationships of care—including caring for children (McLeod et al., 2019), but also healthcare—and knowledge acquisition. In the case of knowledge acquisition, some degree of trust in the source of knowledge is necessary, and John Hardwig (1991) has argued that evidence alone is not sufficient for knowledge, while Mark Own Webb (1993) has expanded on Hardwig’s thesis to defend the necessity of trust not only in scientific knowledge, but in knowledge acquisition more generally.

Finally, a special place in the domain of the instrumental value of trust is occupied by “therapeutic trust,” which refers to the idea that trust can create goods, such as self-respect (McGeer, 2008) and even trustworthiness itself. As McGeer and Pettit (2017) argue, the act of trusting someone has a causal link to their degree of reliability and their desert of trust. In therapeutic trust, trustworthiness is the outcome of trust, rather than its preexisting reason. Being trusted is seen here as a reason to behave in ways that justify trust. In this way, trust also strengthens relationships and forms a virtuous circle of trusting and trustworthy behaviour and expectations.

**III. 2 The Key Ethical Features of Trust**

Three elements of trust can be singled out as posing specific ethical questions: vulnerability, moral obligation, and virtue. We shall examine each in turn.

1. Vulnerability

One of the central reasons why trust poses ethical worries is the vulnerability inherent in any trusting relationship: by trusting, we make ourselves vulnerable, in various ways, to those we trust. A specific way in which this vulnerability is manifested is the possibility for trust to be *betrayed*.[[2]](#footnote-2) In the public domain, the level of trust that the public has in experts and institutions also comes with a corresponding degree of vulnerability. As Katherine Hawley puts it, “trust is needed when we lack knowledge” (Hawley, 2014, p.1), and that is the position in which the general public typically finds itself with regards to experts.

Annette Baier (1986) is the primary source of the argument that vulnerability is a defining element of trust. In her view, trust related vulnerability is explained by the fact that by trusting others, we make ourselves vulnerable to their *goodwill*: “Trust then … is accepted vulnerability to another's possible but not expected ill will (or lack of good will) toward one” (235).[[3]](#footnote-3) We expect a certain degree of goodwill in others, but since their goodwill is something we cannot control, on this account, trust includes giving up some power. Baier shows that there are degrees of vulnerability in trust relationships (ranging from infant-parent relationships to the relative safety of contracts), but that in every case trust impacts power positions. Following Baier, Lawrence Becker (1996) has argued for a similar goodwill-based noncognitive account of trust in political philosophy, where trust is based on “a sense of security about other people’s benevolence, conscientiousness, and reciprocity” (43).[[4]](#footnote-4) Conversely, breaches of trust are harmful, in Becker’s view, not only for individuals but for societies, because our responses to them tend to be more “volatile and disruptive” than responses to mere unreliability or lack or credibility. In this context, the most ethically problematic form of trust identified by Becker is *security*, because lies and mistakes are less likely to erode trust if the trustor construes them as well-meaning, thus introducing a greater degree of malleability in trust, which heightens vulnerability.

The question of the ethics of trust becomes particularly visible when trust is in danger. That is why, as Thomas Simpson (2012) claims, our nuanced discussions of trust arise when trust is breached. If this is true, then a feature of the vulnerability involved in trust is what C. Thi Nguyen (2022) refers to as an “unquestioning attitude.” For Nguyen, trust is neither a belief nor a positive attitude, but a suspension of the deliberative process—such as when we acquire information: “To trust an informational source wholeheartedly is to accept its claims without pausing to worry or evaluate that source’s trustworthiness.” When we trust, we simply do not worry about the reliability of the trustee. This form of trust is a key element of our societies which allows us to outsource much of our deliberation and autonomy. This phenomenon, as Nguyen says, is powerful, necessary, and enormously risky. If trust is unquestioning, the power of the trustee to betray it is much greater.[[5]](#footnote-5) At the same time, this allows us to extend our agency (see also Jones, 2017) and makes our tasks much more efficient, without the need to stop and check every item of information. This brings us to the next point: the desirability of vulnerability in trust.

Vulnerability in trust, including its element of passivity, is not only a cause for worry, but also a positive moral element of ethical relationships. As Baier shows in a later work (2004), an attitude to mutual vulnerability can be virtuous when it contributes to the “climate of trust” (177). Acknowledging our power over one another is fundamental to ethics, and trust creates both power and powerlessness. Partha Dasgupta (1988) and Karen Jones (2004) see the vulnerability of trust as also connected with the passivity of the trustor, which can be virtuous. Jones applies the passivity and vulnerability conditions of trust explicitly to three-place trust (trusting someone in relation to something), defining it as “accepted vulnerability to another person’s power over something one cares about, where (1) the trustor foregoes searching (at the time) for ways to reduce such vulnerability, and (2) the trustor maintains normative expectations of the one-trusted that they not use that power to harm what is entrusted” (Jones, 2004, p. 6).[[6]](#footnote-6)

1. Moral Obligations of Trust

If trust creates vulnerability, then trustees are immediately placed in an ethical relationship with trustors: they are able to respond to that vulnerability either by protecting or exploiting it. For this reason, many have looked at trust from the perspective of the moral responsibility and obligation of the trustee. As Løgstrup claimed, trust makes a “demand” on the trustee (see Stern, 2020). Following Martin Hollis's (1998) distinction between the expectations “that” and expectations “o”’ involved in trust, we do not only expect that experts will tell us the truth: we also expect it *of* them. This seems to place some kind of obligation on the trustee to be trustworthy, but where does that obligation come from?

One common answer is relational: it is the act of trusting that generates responsibility in the trustee to live up to that demand. This type of account depends on the “participant stance” in trust, proposed by Richard Holton (1994). Drawing on Strawson’s reactive attitudes, Holton suggests that trust is one such *attitude*, in which we are not merely observers of another’s actions, but engaged in them, prepared to feel betrayal or gratitude (this also differentiates trust from reliance, in Holton’s view). It should be noted that Holton offers a three-place account of trust, which means that we trust others to do some things, but not others in general (Holton, 1994, p.4). We can adopt a participant stance toward experts in public life, and the three-place specification makes this account applicable, insofar as we trust experts to tell us the truth about scientific information, but not necessarily in other domains. Taking the participant stance means that we hold others *responsible* for what we trust them to do (Hollis, 1998). According to Margaret Urban Walker (2006), the responsibility we ascribe on the trustee comes not only from our expectation that they will act in a certain way, but from our belief that they *ought* to act that way. This way of defining trust, based on what we believe the trustee should do, as we saw in Section III, is known as the “normative expectation theory.” While this theory is dependent on the trustor’s attitude, it is indifferent to the reason why the trustee fulfils the trust-expectation—as long as she does.

For this reason, the normative expectation theory is different from theories that focus more closely on the relationship between trustor and trustee as the source of trust-related obligations, and which ascribe such obligation to the trustee’s awareness of trust being placed on her. These are known as “trust responsive” theories; (see Jones,2013; McGeer & Pettit,2017). The reason for being trustworthy is that one is being counted on: the trustee is *responsive* to the trust placed on her. Note that this does not create an obligation to be trustworthy; it does, however, account for the moral dimension of trust through the responsibility placed by the act of trusting on the trustee, and gives the trustee prima facie reasons to be trustworthy. Also, focusing on the relationship of trust, and on the trustee’s awareness of being trusted, is Russell Hardin’s (2002) suggestion that the trustee is motivated by the thought that the trustor is relying on her, but adding a further normative element: the trustee, when trust is well-placed, “encapsulates” part of the trustor’s interests into her own, and is thus motivated by a combination of inherited interests and the willingness to maintain the relationship with the trustor.

Trust-responsive theories may give the impression that the moral element of trust is primarily second-personal. Paul Faulkner (2014) has reacted against this possibility. According to Faulkner, trust generates not only responsibility but obligation—not, however, because of the demand placed on the trustee by the trustor. Drawing on Løgstrup, Faulkner argues that the fundamental obligation for being trustworthy—here, for telling the truth—is a response to the need of the trustee for trustworthiness. The awareness of trust being placed on the trustee “merely makes [the need] salient” (341). The trustee is bound not only to the trustor, but to her moral commitments to the situation: “What the trusted should do in the trust situation is determined by how things are in the world rather than by the attitudes of the trusting party” (342).

Faulkner’s theory can be placed between two opposed positions, which take trust-related obligation to be based, on the one end, on trust relationships, and on the other, on a commitment to the facts or deeds in question, regardless of whether one is trusted or not. These are the views supported, for example, by Peter Graham (2012), who argues that it is a commitment to truthfulness—rather than to being believed—that is at stake in a trustworthy source, and by Mona Simion (2020), who argues that while speakers do care about being believed, they also, at the same time, care about truth, and that the latter commitment, is more relevant to their trustworthiness than the amount of trust placed in them. Finally, Carolyn McLeod (2002) takes a commitment to her own integrity on the trustee’s part to be the key element to explain the expectation on the part of the trustor.

In criticizing the participant stance, Philip Nickel (2007) and Katherine Hawley (2014) also offer different accounts of the responsibility involved in trust. Hawley, while considering obligation as an important element of trust, concludes that it is too “thin” to be able to define the specificity of the moral dimension of trust: we can fail in our obligations in many ways, without betraying trust (19).I In the sphere of virtue, too, being trustworthy and meeting one’s obligations, again, do not always coincide (11). For these reasons, after considering the partial overlapping of obligation and commitment, Hawley concludes that the responsibility involved in trust is better explained by the idea that “to trust someone to do something is to believe that she has a commitment to doing it, and to rely upon her to meet that commitment” (10). This account is well suited to explain trust in experts, because of their general commitment to the scientific endeavour, and organizations they are part of (see Kelsall, 2021, p. 3). This, also, does not require experts to be committed to the public: differently from Faulkner, Hawley does not take the relationship to be crucial, since the commitment does not have to be to the trustor, or to anything that concerns her, nor does it need to be explicit or motivated by knowledge of the trust placed in them—thus making this account less demanding than preceding ones.

Both Philip Nickel (2007), and Marc Cohen and John Dienhart (2013), instead, take moral obligation to be necessary to a proper understanding of the concept of trust. In both cases, too, the obligation is created by the act of trusting and being trusted. This position is close to the “trust-responsive theories” discussed above, but, at least in Nickel’s formulation, do not require explicit knowledge of trust. In Cohen and Dienhart’s view, in trusting someone, the trustor invites the trustee to accept an obligation; in accepting trust, the trustee takes on an obligation (p. 1). Thus the obligation takes place within a *trust relationship*, and requires both ascription and acceptance. Nickel’s proposal is less restrictive because he does not take the obligation to be actually in place. According to Nickel’s “Obligation-Ascription Thesis,” the trustor *ascribes* and *endorses* an obligation to the trustee (in relation to the object of trust). This, however, does not *create* an obligation. Like Hawley, Nickel does not take obligation to be sufficient to explain trust, but unlike her, he takes it to be conceptually necessary, if we are to make sense of the appropriateness of blame or punishment upon non-performance, and of the motivational force of the ascription. The ascription, moreover, does not need to be explicitly stated, as long as the trustee “reasonably believe[s] the obligation has been ascribed” (Nickel, 2007, p. 317).

The obligation to be trustworthy, by exhibiting the trust-building virtues of openness, honesty and transparency, also depends on the context in which trust is placed. In relation to scientific communication to laypersons, Joshua Kelsall (2021) argues that expert authorities have the obligation to exhibit the above-mentioned virtues, based on the relationships of trust between them and the public. Once again, the obligation is dependent on the relationship, but the obligation is not simply generated by the trust placed on experts. Rather, for Kelsall, the obligation depends on the fact that experts “(1) hold positions of public trust and (2) desire public trust … regardless of whether this builds trust or not” (Kelsall, 2021, p. 288). Legitimate positions of public trust, of course, need to be both deserved and earned: by having the correct qualifications and being able to show them.

Kelsall explicitly rejects Stephen John’s (2018) account, which denies that experts need to have any of the virtues Kelsall calls for, but rather takes it to be sufficient to satisfy to be in a position of trust and to be considered trustworthy. In reality, however, John explains how public trust may depend on false beliefs, which he calls “false folk philosophy of science.” Since parts of the general public may be unable to distinguish between legitimate scientific practices and “folk” scientific practices, John argues that it is permissible for experts to sustain those narratives and thus act against openness and honesty, if that helps to maintain trust that is, in any case, warranted. Kelsall rejects this view as paternalistic and grounded on epistemic interests, which do not lead to the normative claims he endorses. Instead, for Kelsall, “given that the positions of trust often accorded to scientists range over the moral, social, and political domains as well as the epistemic domain, the communicative obligations of scientists must be grounded in concerns about the interests of nonexperts in these further domains” (2021, p. 296). The stress on the fact that scientific expertise inevitably, if sometimes tacitly, brings with it other values, over which scientists have no special claim, is shared by Oughton et al. (2004), who on this basis recommend involving both ethicists and laypeople together with scientific experts when giving advice to policymakers. Moreover, apart from the prudential considerations put forth by John, whether or not the public is able to effectively understand detailed information, honesty and openness are what many want, to the extent that it is part of the first of three demands put forward by the anti-climate change movement Extinction Rebellion.[[7]](#footnote-7)

1. The Virtue of Trustworthiness

So far, we have begun to introduce trustworthiness in connection with the moral responsibility of the trustee. Trustworthiness is considered a virtue. That is why a good amount of the literature on the ethics of trust has looked at what it means to be trustworthy and the conditions under which trustworthiness arises. Baier (2004) includes trustworthiness as a key element of a social “climate of virtue.” Without general trustworthiness, social relationships become strained and difficult.

In the case of experts communicating with the public, as we have seen, trustworthiness tends to include certain other virtues, such as openness, transparency, sincerity, and honesty (see, e.g., Faulkner, 2014 and Wilsdon & Willis, 2004). This is not uncontroversial: according to John (2018), such virtues may not serve the communicative goals well, since distrust of claims—such as those about anthropogenic climate change—are not based on distrust of science per se, but on a poor understanding of how scientific research works. Further, the general public tends to interpret scientific claims as beliefs, whereas a) more often, they communicate the consensus in the scientific community, and b) it is controversial whether groups and organizations can be said to have beliefs (Hawley, 2017). Hence, John concludes, there are other values that weigh more in scientific communication (i.e., furthering non-experts’ epistemic interests) which do not coincide with honesty; and that, given the prevalence of a flawed philosophy of science among the public, transparency and openness may, in fact, undermine public trust.

Another potential challenge to the consensus about the virtues of trustworthy experts is the distinction between “full trustworthiness” and “specific trustworthiness,” drawn by Nancy Nyquist Potter (2002, p. 25). The virtue of trustworthiness, or full trustworthiness, requires being able to be trustworthy overall. Expert trust, on the other hand, typically relies on a “thin” form of trustworthiness, where what matters is experts’ knowledge about the question at hand, and not, say, in their private lives. A further difficulty of the virtue-based account for expertise is that, at least on Potter’s Aristotelian theory of trustworthiness, trust becomes more demanding: in Potter’s view, it is not sufficient that one proves to be trustworthy on particular occasions; trustworthiness is a character trait, and as such, it a) needs to be developed over time, and b) needs to adhere to the “mean,” requiring practical wisdom for its right application in the right situations for the right end. This does not make the theory inapplicable. Two features open it up for experts. The first is the contextual nature of the theory. On a virtue-based account, trustworthiness needs to be appropriate to the situation, and public trust is different from individual trust. Secondly, experts can indeed embody the dispositional qualities that Potter suggests: giving assurance and signs of their trustworthiness; taking their responsibility as knower seriously; and responding appropriately to broken trust and taking steps to mend it.

All this suggests that we can take some elements of the virtue-based account of trustworthiness to examine trust in experts, but its applicability is limited. As Karen Jones (2013) argues, the virtue account does not make sense of conflicting trust requirements, nor does it explain why being untrustworthy is not always a vice. However, not everyone agrees that individual relationships of trust are radically different from the public trust relationships between experts and the public. Arguing for the importance of trust in climate change communication, Ben Almassi (2012) claims that it is possible and indeed necessary to build “mutual and reciprocal affective trustworthiness … to avoid moral rot” (p. 30). On this model, trustworthiness needs to go beyond reliability and transparency, to include the recognition, in both parties, of the epistemic dependency of the public, and a “conscientious response” on the part of experts to this dependency.

**III. 4 Reflections on the Ethical Responsibility to Trust**

Much of the interest in the ethics of trust concerns the ethical responsibility to be trustworthy and the ethical responsibility to establish conditions of trust. But we also want to consider the less theorized issue of the ethical responsibility to trust or distrust. Even where the goal is to establish social conditions in which people end up trusting trustworthy sources, one may get a better sense of these conditions by treating the people who we want to be trusting appropriately as moral agents—agents who are trusting or not trusting in response to what they take to be moral imperatives and not merely psychological cues.

This means we treat people who distrust reputable journalists and scientific experts and instead trust populist politicians and internet conspiracy theorists as rational agents (but see Boghossian in this volume), responding to genuinely ethical perspectives, rather than as psychologically twisted and lacking fully autonomous agency. This enables us to see such people from their own point of view—with empathy—and to engage with their perspective. The advantage of this approach is not just that it is more respectful and humble, but also that it opens up the possibility of genuine dialogue.

The issues involved are complex and are ofthe underpinned by implicit or explicit theoretical structures for thinking about trust. Roughly speaking, there are two main theoretical approaches. In one, represented recently by the work of Paul Faulkner (2015) and Jacobo Domenicucci and Richard Holton (2017), trust is treated in the first instance as a two-place relation between people. It is taken to be an attitude parallel to the attitude of loving someone—an attitude that may be understood without reference to any specific action it disposes one to perform. It disposes one to do nothing more specific than to grant discretion to the other person, and what this entails will vary from situation to situation.[[8]](#footnote-8)

According to this approach, it is possible to understand what it means for someone to trust another person without relating this to any disposition to trust them to do something in particular. Thus, understanding what it is for someone to do something in particular—the three-place relation of trust—should be based on an understanding of this general two-place relation of trust.

This approach faces some obvious difficulties. The most obvious, perhaps, is that a person may trust somebody to do one thing but not trust them to do another. However much I trust my friend in many ways—and trust them in particular to stand up for me if I need moral support—I might not trust them to bail me out of financial difficulties or save me from an escaped lion.

The other theoretical approach to trust takes the trusting *behavior,* rather than the trusting *attitude,* to be basic. So the three-place relation of a person trusting another person to do something is taken to be more fundamental in the order of explanation than the two-place relation of a person trusting another person. As Harold Noonan (2021) has argued, the two-place relation may be understood in terms of the three-place relation as follows: A trusts B if for a variety of tasks fixed by the context, A trusts B to be doing those tasks when required. Indeed, it is really a four-place relation—A, in doing φ, trusts B to be doing ψ—that should be treated as basic.

What is key to that relation is the idea of reliance. A, in doing φ*,* is doing something with a goal that requires B to do their thing—ψ. And an extra condition that is essential to this four-place relation, but is not always brought out in the philosophical analyses, is that what A is doing and what they are relying on B to be doing are part of the same plan or enterprise. The behavior that trust belongs in is collaboration. If A is trusting B, then they are doing something as part of a collaborative enterprise, where, by doing it, they are relying on B doing their part in the same enterprise. To take an example, in a game of football, a successful pass would depend on the activity of two people—the person kicking the ball and the person running into space to receive the pass. So, a pass is a collaboration. If one player runs into space as part of that collaboration trusting the other to kick the ball through the channel, they are trusting their teammate to be doing that. At the same time, the player kicking the ball is trusting their teammate to be making the run they had anticipated.

Consider another example: getting vaccinated. When I receive the jab, I trust the medical experts to be supplying something safe and effective. The collaborative enterprise we are both engaged in is the vaccination program. In doing my part I am relying on them to be doing their part. It is not quite true that they are also relying on me to do my part, though they are relying on enough of us to be doing our parts. So, we might say that the medical experts are trusting the public at large to do their part in the vaccination program.

This simple idea that trust involves reliance within collaborations is sufficient to generate some interesting analysis of the ethical responsibilities to trust. The first question it raises in this context is what sorts of collaborative enterprises might generate ethical responsibilities. In the context of the PERITIA project, it is the collaborative enterprise of knowledge transfer. The ethical responsibilities generated by this enterprise alone are quite thin but remain of interest.

Knowledge transfer is a collaborative enterprise involving someone who knows and someone who learns. They are both essential to the process. The part of the expert—the knower—is to communicate truly and helpfully, and the part of the learner is to understand and accept this, or in other words, want to learn.

Looking at it this way, it becomes clear that two things are required for someone to be motivated to accept what someone else is telling them. First, they must be motivated to engage in the collaborative enterprise of learning. Second, they must be motivated to rely on that particular piece of testimony—that is, they must be willing to rely on the expert doing their part, namely, telling the truth.

Ethical responsibility for the receiver of the knowledge only enters into this picture to the extent to which that individual has the ethical responsibility to know. While knowledge may be an end in itself, there is no ethical responsibility to know everything. It is no failing of mine not to be interested in knowing which teams have won the men’s football World Cup over the last sixty years. And so the ethical issue of trust or distrust does not arise for me if some football fanatic is telling me the list of winners. On the other hand, I should know the names of the streets in my town if I am presenting myself as a taxi driver. As an academic, I should know the regulations that govern my activities, and I should know a good deal about the subject I am teaching. Arguably, we all also have an ethical responsibility to know what is going on in the world to the extent that we are all potentially engaged with the world and with one another in social contexts.

So the ethical responsibility to know things within some realm depends on one’s social role with respect to that realm. And when you do have an ethical responsibility to know something which you cannot work out for yourself, then you have a corresponding responsibility to engage in a collaborative exercise of knowledge transfer and reception. This means we have an ethical responsibility to trust or distrust appropriately։ to accept what a trustworthy person tells us, and not what an untrustworthy person imparts, or at least, to do our best to distinguish between the two. This is an ethical “should.” And if we do not take sufficient care to establish whom to trust—at least on occasions where we have an ethical responsibility to know—then that lack of care is ethically blameworthy epistemic negligence.

The prime minister who is imposing a new set of rules on the citizens of a country concerning the banning of parties during a pandemic is under an ethical obligation to know what those rules are. And if s/he trusts the word of a senior civil servant who says it will be fine to have lots of drinks, karaoke, and party games until two in the morning because the event takes place in one’s workplace, then that trust might be regarded as ethically negligent.

**Part IV. Dealing with Distrust**

Before examining the question of distrust in science, we would first like to examine the question what it means to trust science and scientists. One difficulty is the applicability of reflections on trust, often generated in relation to individuals, to the case of science and scientific experts. Scientific policy advice is generally produced by groups rather than individuals. This, on the one hand, can lead to greater trustworthiness, as Naomi Oreskes (2019) has argued, because the collective nature of science means greater checks and lower fallibility. On the other hand, while it is intuitively easy to think about trust in individuals, trust in groups may seem too abstract to grasp.[[9]](#footnote-9) The absence of some elements of trust such as mutual relationships, personal acquaintance, and emotional affect seems to suggest that trust, due to its ethical dimension, is not a fitting concept for science and the detached objectivity that science aims to achieve. Arguing against those who believe that objectivity requires a shared basis for trust, and that trust is necessary on the part of the public to believe experts (e.g., Scheman, 2011), Koskinen (2020) claims that they fail to distinguish between trust and reliance, and that the latter is more applicable to scientific objectivity: following Baier (1986), it is trust that can be betrayed, not reliance; people can betray us, not processes. Koskinen denies that we can trust science, but not that we can trust scientists. However, she introduces a difficulty insofar as the object of trust in scientists is science, and scientists are representative of scientific objectivity.

Using the same distinction, Hawley (2017) argues that, with groups, trust and reliance are far more similar than they are in the case of individuals. Hawley compares trust and reliability in individuals and in groups, focusing on testimony, and concludes that in group testimony, the practices can be described without resorting to the distinction between trust and reliance. A further difficulty with the very question of the difference between individual and group trust is the fact that in public communication it is not straightforward to make this distinction in the first place. Jennifer Lackey (2018) argues that, when a group “speaks through” a spokesperson, the group’s testimony simply is the spokesperson’s testimony; for John (2018), individual scientists represent the consensus; while Hawley (2017) claims that responsibility for assertion lies with the group. Further, Hawley suggests that in many cases of trust in groups—such as scientists—the public is actually trusting (or distrusting) individual scientists; conversely, group membership influences the level of trust in individuals. This makes identifying the specificity of trust in experts as a group rather difficult.

Nonetheless, as Hawley points out, in ordinary language we do talk meaningfully about trust in entities other than individuals, including groups, and at least some features of individual trust, as we have seen, remain applicable to group trust. Whether public trust is more properly applicable to the individual scientist or the group, and whether the spokesperson is properly trusted as an individual or representative of group or group consensus, it remains true that the public has attitudes, beliefs, and emotions of trust and distrust toward scientific experts, and that this connects with the ethical issues of public’s vulnerability and autonomy, expert’s authority and power, and the practical beneficial or detrimental consequences of these trust relationships for the general public.

The substitution of trust by reliance, moreover, does not work if we consider the situational demands of specific cases of the science/policy nexus. To take an example for our recent shared past, trust, rather than mere reliance, often was the appropriate doxastic attitude toward the COVID-19 stringent public policy measures and calls for vaccination. The conditions of uncertainty and novelty we faced frequently involved the acceptance of risk in the face of the possibility of harm and substantial sacrifice of bodily autonomy by following expert advice. Unlike reliance, trust invokes a range of emotional and evaluative responses, including a sense of risk-taking, the possibility of feeling betrayed (Baier, 1986), and an engaged attitude rather than a spectatorial one (Holton, 1994). As we have seen, when we trust, we make ourselves vulnerable, and this is especially true when trust requires us to *act* in specific ways, and these actions could affect our lifestyle or our health. The trust required, of course, is not blind, and is informed, among other things, by the reputation, the track record of past performances, and the success or failures of the experts. But in the case of COVID-19, the extent of the knowledge gap between the experts and laypeople, coupled with the urgency of decisions based on scientific advice and the high stakes involved, called for the riskier attitude of trust rather than mere informed reliance. And yet, trust in experts poses its own normative challenges which can translate into lack of trust or increased distrust. There are, however, good and bad reasons for distrust of experts. In what follows, we address and dismiss some of the more spurious reasons for the distrust of experts before discussing some legitimate social and ethical concerns that lead to such distrust.

1. Loss of Epistemic Autonomy

The unavoidable lack of knowledge in specific domains in the general public makes it necessary to acquire it from experts in a number of situations. But this necessity seems to threaten a key principle of post-Enlightenment Western morality, and the key principle of Kantian ethics, namely autonomy. Applied to knowledge, this leads to epistemological individualism, or “think for yourself.” However, as Koskinen (2020) points out, scientific practices requires that experts, too, trust other experts in their knowledge acquisition. In an age of hyper-specialization, as Elijah Millgram (2015) argues, the ideal of complete intellectual autonomy is even more distant. Epistemic dependence, and the consequent threat to autonomy, are not only a problem for the “general public.” Such a position is unavoidable, and the alternative appears much worse, bringing about, as Hardwig (1985, p. 343) puts it, “a situation in which he must either suspend belief or—if this is impossible or undesirable—arrive at belief on some admittedly nonrational basis.”

Rather than abandoning autonomy, this seems to point to a flawed conception of autonomy, one that, incidentally, is also influenced by social structures and status. Grasswick (2014) offers evidence of this in her study on trust and climate change. Drawing on feminist epistemology’s situated approach, she shows how trust in institutions of climate change science also depends on the relative privilege of the trustor. Grasswick concludes that greater privilege equals diminished trust, with the case of white males being more likely to distrust expert information about climate change. Individuals from privileged backgrounds appear more confident in their own “autonomy” and less likely to rely on expert information even when there are good reasons to do so.

There are rational—and Kantian—reasons to subscribe to a modified conception of autonomy that makes room for trust. As Onora O’Neill argues, autonomy does not need to extend to personal whim (as in J.S. Mill’s model) but can be guided by practical reasons (as in Kant’s model). Writing in the context of bioethics, where autonomy is considered to be one of four ethical “pillars,” O’Neill shows how Kantian autonomy can in fact be the basis for fostering trustworthiness.[[10]](#footnote-10) Following O’Neill’s strategy of distinguishing types of autonomy, Nguyen (2018) argues (against Millgram’s more pessimistic view) that we can preserve autonomy in trust. The kind of intellectual autonomy that is threatened by trust in experts, according to Nguyen, is the more common type, *direct autonomy*, where we seek to understand the information and process by ourselves. This is the kind of autonomy that requires transparency. However, there are two other kinds of autonomy, which are consistent with trust but not with transparency: *delegational autonomy* and *management autonomy*. In the former, we trust others to provide information we cannot arrive at ourselves, but remain autonomous insofar as our act of trust is active and justified—we take responsibility for it, and for our choice to delegate; in the latter, we put together, for ourselves, information from different sources and take responsibility for the whole—the larger system of knowledge.

1. The Democratic Deficit of Reliance on Experts

The loss of private epistemic autonomy also has a public counterpart.  *As we have seen in Section 1,* experts’ involvement in policy advice has become a constant feature of modern governance. For a time, *epistocracies* were regarded as harbingers of a new political era, in which ideologies would give way to technical problem-solving and expert-driven, problem-oriented thinking would replace both capitalist and socialist states (e.g., Price, 1965). While these high hopes proved ill-founded, reliance on experts in policy decisions in a wide range of areas, from economics to health, from technological know-how to agriculture, remains central to contemporary governance. The so-called “knowledge economies”—where intellectual capital is a key source of economic growth—accentuates this reliance.

There are two interconnected concerns specifically about the role of experts in democratic governance: first, expert knowledge may not be readily open to assessment by non-experts, and therefore, the very idea of expertise is premised on epistemic inequality which has social and political consequences. Second, experts, when they function as unelected contributors to political governance, are immune from the type of accountability we impose on elected members of governments. Experts, by definition, know more than the general public about their area of expertise, and particularly in knowledge economies those with greater epistemic resources and access to such resources frequently—but by no means invariably—are rewarded financially and accrue prestige and status, becoming the new elites. The top-down reliance on a small group of experts on policy decisions leaves less space for the participation of citizenry in political decision making. Status inequality, Hugh Desmond (2022) has argued, can be a greater reason for distrust than other epistemic or moral reasons.

This gives rise to democratic deficit of trust in experts: if policy decisions are to be taken based on advice from experts, and if the correct epistemic position toward experts is one of trust rather than critical questioning, then the space available not just for political contestation but also for political decision making will shrink in proportion to the extent that we allow expert-driven trust-based policies to become the guiding principle of political governance. These worries offer grounds for scepticism—and indeed distrust—of experts, not primarily because the content and sources of their advice, but because of the impact of their involvement in political decision-making. Yet, there is no easy distinction between these two reasons: in promoting trust in experts, the question of its short- and long-term impact is pertinent to our judgements of trust.[[11]](#footnote-11)

**VII. 1 Decreasing the Democratic Deficit of Trust in Experts**

There are, however, social tools at our disposal to address the issue. Citizen participation, at various levels of knowledge creation and transmission (or what Dewey calls “popularization of knowledge”), is a good example. The point is echoed by many contemporary thinkers. Philip Kitcher (2011), for instance, argues that to counter value-based distrust, what is needed is a more inclusive participation of the public in the workings of scientific research from the very beginning, including, in particular, the facilitation of a more transparent and open discussion about the values that inform such research.[[12]](#footnote-12)

An effective way of including the public in expert-informed policymaking is the routine use of citizen assemblies and mini fora (Farrell & Suiter, 2019) where experts and representatives of the general public engage in publicly accessible conversations and deliberations about the policy choices available. Making intellectual space for the contestation of dominant views, and monitoring the performance of experts and their commitment to honesty, transparency, and good will are further means of both creating a climate of trust and enabling a participatory form of democracy in which experts and policymakers are held accountable to the citizenry.[[13]](#footnote-13)

A division of epistemic labor is essential for the smooth running of any society, but the division does not need to be purely hierarchical. A horizontal model, where multidisciplinary panels, including lay members of society share the responsibility for policy advice, is more in line with the egalitarian aspirations of democratic governance. Such “horizontal” models do not deny the role and significance of specialists’ knowledge in decision-making, nor do they flatten the idea of expertise by placing their level of knowledge and on par with the laypeople’s. Rather, they allow for input from and debate between a variety of sources and voices. The complaints about the elitism of experts, more often than not, are directed at the exorbitant financial rewards they receive, as well as the air of arrogance surrounding them, rather than the knowledge and information that they possess. The horizontal conception of the division of epistemic labour on expert panels can also have beneficial consequences for structuring the financial rewards that experts receive and thus address some of the grave economic inequalities that the knowledge economy has engendered. The horizontal model, put into action through citizens’ assemblies, is part of a move toward a more cooperative rather than hierarchical division of epistemic labour, which also constitutes the reciprocal nature of the demands and commitments that are part of a climate of trust. Efforts to make the practice of epistemic humility more commonplace among such experts can also be helpful.

1. Trust, Distrust, and Social Identity

The possibility of retaining autonomy in trusting experts does not remove the danger of not being able to do so, nor does it remove the inherent vulnerability and power structures of trust. This is why, as Jason D’Cruz (2019) has argued, both trust and distrust carry significant moral problems. The power relations underlying trust in experts are not only of an epistemological kind. Sometimes, experts are trusted not because of the power they acquire through greater knowledge, but for their political and social power. This is an instance of poorly justified trust. Conversely, oppressed groups which face discrimination in other contexts of life, also tend to face epistemic discrimination, or “epistemic injustice” (Fricker, 2006). This occurs when their testimony is treated as less relevant or competent for reasons that are not grounded in good epistemic practices, but in social and political bias and prejudice. This is an increasingly studied phenomenon in moral epistemology, especially feminist epistemology (see Code, 1991, especially Chapter 5), where it is argued that not only are women experts less likely to be taken seriously, but that the very model of knowledge is based on patriarchal structures. Like Code, Linda Martin Alcoff (1999) argues that social identity is relevant to epistemic credibility. Insisting that epistemic practices are and should be “neutral” is therefore inadequate. These phenomena have parallel repercussions in the case of trust. A female scientist or one from an outgroup, for instance, may find it more difficult to garner trust in her work than a white male scientist from an elite educational background. The well-known story of the eventual Nobel prize winner Katalin Karikó’s struggles to gain recognition for her pioneering work on mRNA vaccine is a case in point.

Kariko’s case is a pointer to the relevance of the social identity of the trustor and trustee. As Hawley (2017) points out, both trusting and distrusting are exercises of social power, where what is at stake is both the social identity of the trustor and the trustee—the latter influencing the very possibility of her becoming an object of trust at all.

One kind of testimonial injustice, as identified by Fricker (2006) is treating someone as a mere source of information, rather than as an informant (who is part of a community with shared goals), resulting in objectification. Hawley (2017) applies this point to trust, suggesting that objectification occurs when we offer people mere reliance in situations where the richer attitude of trust would be appropriate; denying people trust is a form of disrespect. Hawley concludes that the form of injustice involved in lack of trust cannot entirely be explained on epistemic grounds: sometimes it is epistemically permissible to be more and less trusting, but “it is not morally permissible to switch between these doxastic policies on grounds of social identity” (Hawley, 2017, p. 78). More broadly, as Govier (1992) has pointed out, if trust requires some degree of collaboration or partnership, inequality will get in the way of collaboration and hence of trust, so that these matters need to be addressed, individually and institutionally, for minimizing the influence that a history of oppression has on relationships of trust. Naomi Scheman (2011) has shown how bias and injustice can penetrate very deeply into the discourse of science, so that the very appeal to objectivity can entail a removal of social perspectives, which block some participants from entering the discourse of objectivity. Only once this is acknowledged can the trustworthiness of science be convincingly demonstrated.

Of particular concern is when trust lies mostly on the part of—as well as being encouraged from—vulnerable groups, including social and physical vulnerability. That does not only add to the inherent vulnerability of trust relationships, but also reinforces pre-existing power structures of political and epistemic inequality. Catriona Mackenzie (2014, 2020) has distinguished between inherent, situational, and pathogenic vulnerabilities. Inherent vulnerability, embedded in the human condition and situational vulnerability, which arises from particular contexts, are exacerbated by pathogenic vulnerability—where we are simultaneously more in need of trust and more at risk because of it. These cases highlight how trust can at the same time open up valuable opportunities and put us at great risk. Drawing on feminist literature, Anita Ho (2011) examines the case of trust between healthcare providers and people with impairments, arguing that the burden and vulnerability of trust lies excessively on the latter’s side, perpetuating social injustice. Their condition makes people with impairments less likely to be trusted and heightens the epistemic and power superiority of the healthcare providers in ways that are also not conducive to the best treatment. Ho calls for greater collaboration between professionals and people with impairments, a greater humility in professionals, and a greater appreciation of the knowledge and experience that people with impairments can bring to the conversation about healthcare.

1. *Concerns Arising from the Value-ladenness of Science*

Science sceptics claim, on a variety of grounds, that science does not deliver the objective, interest-free knowledge it promises. Such scepticism may be directed at the corrupt practices of individual scientists, or it may be seen as inherent in the very methodology of science. Both types of criticism result in the accusation that the scientific theories and their resultant technical knowledge that inform policy decisions are never pure or value neutral, but to the contrary, they are often infused with personal and ideological biases that support the interests of the individual scientists and/or the existing economic and power hierarchies. We will look at each concern in turn.[[14]](#footnote-14)

The simplest reason for distrust of experts is the suspicion that their advice is informed by their personal and sectoral biases or financial interests. Examples of fraud, personal bias, and incompetence, while not very common, are part of the landscape of expertise. Undoubtedly, there are bad actors among experts whose advice is motivated by personal or professional gain rather than the best scientific evidence, but moral or professional failures at this personal level are not a very good reason for scepticism about the science/policy nexus. While blatant corruption is a sad social reality, there are reliable ways of detecting and addressing such corruption in scientific practice—for instance, through peer review and expectations of the replicability of results, particularly where the results are surprising or indicate significant breakthroughs; whistle-blowers have also played an important role in exposing corrupt practices at institutional level. For these reasons, fraudulent practices within science are not very common.

There is a host of deeper and subtler reasons for doubts about the purity and value neutrality of scientific advice—reasons that are embedded within the operational and theoretical frameworks of science rather than in the psychology or the proclivities of individual scientists. First, there are concerns about the underdetermination of scientific hypotheses by existing data, i.e., that there are empirically equivalent rival theories which are equally adequate in explaining experimental results or observations (Quine, 1970, p. 179). Underdetermination poses the question of how scientists choose between different empirically adequate rivals and the extent to which values play a role in such decisions. The so-called Problem of Unconceived Alternatives, based on the historical evidence that “typically [there] are alternatives to our best theories equally well confirmed by the evidence, even when we are unable to conceive of them at the time” (Stanford, 2001, S9), shines a different light on the same problem. The “new pessimistic meta induction” from past failures in imagining better theoretical alternatives opens the possibility that we may be in the grip of similar failures now and that our biases and values may not allow us to imagine and give preference to alternative scientific theories. Finally, the Inductive risk argument (Hempel, 1965; Douglas, 2013), starts with the observation that scientists never have conclusive proof for their theories or complete evidence for their hypotheses, but always face a degree of uncertainty regarding scientific knowledge, so they need to use their judgement to make a final call on how much uncertainty they are going to accept. Such judgements always have strong normative elements.

What the three arguments show is that theory choice is not fully determined by available evidence: scientists use their judgement, exercise imagination, and make risk assessments in prioritizing a particular theory over and above others, and in doing so, they inevitably rely on value judgements. Moreover, contra Kuhn (1977), such judgements are not restricted to epistemic values only. As Heather Douglas (2013) has argued, values—both epistemic and moral—come to play an important role in framing the problems scientists address, deciding on the range of evidence they consider, the scope they assign to a hypothesis, the level of uncertainty they are willing to accept, and how they calculate the consequences of potential error, among other factors. The optimists about science, and Douglas is among them, believe that scientists should, and are in a position to, bring in values such as benevolence (or a concern for the welfare of others) and the principles of least harm and due diligence in assessing negative consequences to ensure that the value gap in science is addressed in ways that are beneficial to the recipients of their advice. Pessimists, among them some feminist epistemologists, on the other hand, point out that scientists’ theoretical choices are frequently coloured by unacknowledged gender, race, class, and other ideological biases. Moreover, in most cases, only a historic distance will allow us to detect the full range of pernicious values that are brought to bear on the choices of theory and evidence. The upshot is that, one way or another, theory choice is not value-natural, and the values that are brought to bear have a social and moral dimension.

1. *Accommodating Values*

Trustworthy science attempts to acknowledge the role of values in sciences and to come to terms with it in appropriate ways. The first step toward such accommodation is to acknowledge the uncertainties of science and the value gap between evidence and theory, ensuring in particular that the gaps are filled in such a way that the well-being and best interest of those affected by the activities and findings of the scientists concerned become central to the conduct of science. This transparency and acknowledgment of value would go some way toward addressing the concerns expressed above.

The second step is the acknowledgment of the values that may inform the public’s inclination or disinclination to trust specific scientific information. The creation of a climate of trust requires an awareness that distrust in scientists may stem from genuine differences between the public and experts regarding the values that shape scientific research (Kitcher, 2011). In what is presented as a purely scientific question, different values, priorities, and preferences are involved, and bringing them out may lead to a more respectful, but also more mutually trusting practice. So, Popa (2024) argues for assessing trust-conducive values in public health. Also helpful is addressing the lack of clear diversity and broad representation in the scientific community, as well as the insufficient acknowledgement of the plurality of values that determine trust and distrust in science. To take one example, climate sceptics appear highly susceptible to messages delivered by their social group, particularly their own political group. This demonstrates not only reluctance to change, but also the importance of group belonging when it comes to trust. Group belonging here includes values: shared conservative values have been shown to play an important role in whom to trust. For these reasons, David Hall (2019) has suggested that in order to address climate scepticism, it is necessary to “frame” the message without denying or eradicating the fundamental values of some groups. Drawing on Bernard Williams’s (1979) “internal reasons,” Hall suggests a model of persuasion which is truthful, and also which acknowledges the plurality of people’s motivations, connecting facts about climate change with people’s *subjective motivational sets* (Hall, 2019, pp. 41-42). Similarly, Feygina, Jost, and Goldsmith (2010) suggest that reframing pro-environmental change as, for instance, preserving, rather than challenging, the social system (e.g., the “American way of life”) may encourage those who are motivated to protect the system to take greater personal responsibility (p. 333, see also Kahan, 2010).[[15]](#footnote-15)

1. *The Place of Distrust*

So far, we have discussed distrust as a problem, mostly a result of inequality or oppression, with worrying implications for trust in experts. Trust, furthermore, is a form of recognition and respect, and misplaced distrust comes at epistemic, social, and moral costs (see Jones, 2013). But there are cases where distrust is not only epistemically, but also morally justified.[[16]](#footnote-16)

Distrust may be well placed exactly when trust is demanded vocally and with great insistence—there is something suspicious about the desire to be trusted. This is corroborated by the fact that, as Becker has observed in the case of political trust,

Modern dictators … have often exhibited obsessive concern for maintaining the credulity and dependence of their citizens and civil servants… By contrast, the ruling elites in liberal democracies and open-market economies seem able to tolerate a great deal of incredulity in their citizenries (Becker, 1996, p. 60).

This seems to suggest, to generalize, that the greater the demand for trust—especially when it is not accompanied by trust in the other party—the more justified the distrust. There are also cases in which trust is forced, or is less freely given than in other situations. This could be applied to the question of climate change, at least whenever the public is convinced of the urgency and seriousness of the threat.

Meena Krishnamurthy has argued that distrust not only can be justified, but *valuable* in a democracy. Taking as her example the Black Civil Rights movement in the United States, she suggests that it was distrust that was central to Martin Luther Kings’ actions and the building of a less oppressive system. The value of distrust, Krishnamurthy argues, lies in its “tendency to bring about justice by tempering tyranny” (Krishnamurthy, 2015, p. 400). In the case she examines, distrust led to more careful but also more forceful and creative action, as well as the creation of a system of checks and balances (through the Constitution and Bill of Rights) to prevent abuses of power.

Within democratic societies, distrust may arise appropriately in response to power relationships, but also in response to disagreements about values. As Philip Kitcher (2011) argues, distrust in scientists may be the result of a genuine difference between the public and the experts about the values that shaped scientific research. Similarly, Maya Goldenberg (2016) refuses to take a paternalistic approach to the question of trust in experts in the context of vaccine hesitancy, arguing that hesitancy may not be driven solely by ignorance and stubbornness, but grounded in the failure of experts to address parents’ concerns about their children’s health. Once again, in what is presented as a purely scientific question, different values, priorities, and preferences are involved, and bringing them out may lead to a more respectful, but also more mutually trusting practice. For both Kitcher and Goldenberg, the fact that this type of value-based distrust exists calls for a more inclusive participation of the public in the workings of scientific research from the very outset, including in a more transparent and open discussion of which values inform such research.

The strategy to build trust—including “social trust”—based on transparency and collaboration, through collective reasoning about norms and how they are developed, has also been suggested in situations of oppression, e.g., by Shay Welch (2013). In some cases, however, these conversations about values aimed at restoring trust cannot take place because the moral positions involved are incommensurable. This is the case of distrust in science when it is based, or has historically been based, on sexism or racism (e.g., the Tuskegee Experiment, see Thomas &Quinn, 1991).

Exploring a rather under-discussed aspect of trust, Katherine Furman (2020) looks at the affective dimension of trust and distrust in science. She argues that epistemic and value-based explanations of distrust are important but insufficient, and that emotions often play an important role in understanding the dynamics of distrust. Exploring cases of medical abuse in South and West Africa, Furman shows how emotions accompany distrust (fear, frustration) and pre-date it and exacerbate it (the anxiety associated with poverty and discrimination), explaining the disempowerment that comes with it. Furman also offers a suggestion for how to overcome distrust in cases such as these, where there are good political and historical reasons for distrust in experts, but where such distrust is detrimental to the agents themselves. Two key asymmetries—power and cost—need to be addressed. Based on the example of the doctor from Doctors without Borders (narrated in Steinberg, 2017) in South Africa, who jumped on a table and drew his own blood to show that the HIV tests he was offering were safe, Furman shows how power and cost asymmetries can be diminished, so that the relative affective component is also rebalanced.

Distrust, like trust, can have deeper roots than it first appears, and may be grounded in social and power dynamics that fundamentally determine the way expert information is shaped and received. Distrust, in these cases, is not necessarily a problem; in fact, it can point toward inequalities that influence epistemic practices, and that need to be rectified. For this reason, we have argued, in cases where trust in experts is urgently called for, it is not sufficient to clarify information, share more data, and demonstrate experts’ honesty and reliability. What is needed, we suggest, is a larger scale intervention to foster what Annette Baeir (1986) has called a “climate of trust,” in which the costs and vulnerabilities of trust are shared, and which enables and encourages the participation of various groups, with differing values, both in the production of knowledge and in the outcomes of that knowledge which—in the cases examined—have significant impacts on everyone involved.

**Conclusion: Toward a Climate of Trust**

The concerns regarding trust in experts and its ethical implications outlined above show how the question of trust in expert advice could be seen as a legitimate terrain of both normative and political contestation. To withhold trust in experts, particularly in the political domain, can be, and often is, a political act. However, it should not be dismissed merely as a by-product of right-wing populism, even if in many recent instances it has taken that form. The cumulative impact of the above considerations (a-e) is to demonstrate that there are legitimate grounds for scepticism toward experts, particularly in the context of the roles they assume in public life.

Reasoned scepticism about the role of science in policy decisions should be taken seriously, both because of the worries about the politics of trust raised above, and because distrust may *signal* social imbalances and wrongs in the way scientific knowledge is placed in the service of the political and social interests of some sectors of the society at the expense of others. In such instances, scepticism and distrust are not only justified, but may have the value of pointing in the direction of something that needs rectifying. At other times, distrust stems from a fear of losing autonomy and, hence, power. This source of distrust is different, but it too points at social imbalances that need to be acknowledged. On the other hand, as we have seen, distrust of experts, even when it originates from experiences of injustice and long-standing grievances, becomes a source of concern when scientific advice plays a crucial and urgent role in securing the wellbeing of a population or a planet. As the COVID-19 pandemic has dramatically demonstrated, expert intervention in policy decisions can be of critical urgency, and the need for their advice, literally, a matter of life and death. This point can be generalized beyond the current emergency and be applied to other threats facing us, with climate change and vaccination being among the most notable examples.

In such instances, trust in experts takes on specific features, which are relevant to assessing its ethical significance. First, health emergencies or the science of climate change are backed by widespread consensus within the scientific community, offering grounds for reduced scepticism. Second, the urgent need for action in the face of an imminent threat reduces the opportunity for independent reflection and assessment, and heightens the need to exercise trust as a shortcut to immediate action. Third, trust is not just expedient but essential in these matters, because firsthand knowledge of medical and climate science among the public is scarce, and even well-informed attitudes about these questions essentially depend on trust in the testimony of experts (see Almassi, 2012). Furthermore, the requirements of compliance with expert advice for the public good in these cases is not just a question of individual autonomous choice, but a matter of significance to all members of society, in the local as well as global community. The prominence of the cases of the COVID-19 emergency and climate change in the discussion of trust is motivated by the urgency of acting upon the information conveyed, and the far-reaching consequences of either trust or distrust in relevant expert advice.

Much of the reaction to the real or perceived breakdown of trust in experts, over the last few years, has called on better communication or messaging by scientists and science journalists (Lewenstein & Brossard, 2006; Jasanoff, 2014), greater scientific literacy on part of the general public (Lombrozo et al., 2008; Miller, 1983; Miller, 2004; Bodmer, 1985), and countering the impact of motivated cognition (Oberauer & Lewandowsky, 2016). It is assumed that addressing these flaws will also address issues of distrust in science. Others have rightly emphasized the need not only to substantiate but also to increase the trustworthiness of experts by demonstrating not just their knowledge and competence, but also their honesty in communication and their responsibility or responsiveness to the evidence (Anderson, 2011, pp. 145-46). Nevertheless, as we have seen, distrust is not occasioned solely by perceptions of incompetence or failures in performance. Nor is it only the result of normative and ethical failures, such as dishonesty or irresponsibility, on the part of experts.

Distrust arising from socially based concerns demonstrates that what is at stake is not just professional credibility or competence, nor intellectual desire to master more information. In fact, distrust in experts in these cases can arise from the *rejection* of the information coming from a particular source, not just because of worries about the accuracy of the information or the personal credibility of the experts responsible, but because of the background values and political structures within which the information is created and shared. What is at issue is not only the *content* of the message of distrust or scepticism, but also both the *identity* of those who are distrustful,[[17]](#footnote-17) as well as the social, political, and historical factors influencing the creation and reception of the scientific message.

To reiterate the point, while ensuring the trustworthiness of experts—and of the policymakers they advise—is essential for countering unwarranted distrust, we also need to take into account the varying factors that go into the legitimization of distrust. To achieve the complex goal of countering the call of distrust, we argue, it is necessary to create a climate of trust—a social and political environment in which the concerns that legitimize distrust are acknowledged and, insofar as possible, addressed, and where legitimate trust is allowed to flourish.

Trust in experts, particularly but not only in cases where they guide policy decision, cannot be treated in isolation, for it is interwoven with other social and political forms and requirements of trust. For this reason, as Scheman (2011) has argued, the impersonal discourse of science, its universal concept of objectivity, and the impersonal demonstrability of trustworthiness that goes with it, are not enough in a context where social hierarchies exclude some individuals and groups from the production of knowledge and the exercise of power. Trust, as Scheman puts it, “needs to be convincingly demonstrated—not just abstractly demonstrable,” and justified belief in the trustworthiness of the scientific methods practiced by institutions depends on:

the justified belief that those institutions do in practice what they are supposed to do in theory: ground knowledge claims that are acceptable to all of us, not just to those of us with certain forms of privilege, who see the world through certain lenses, from certain biased perspectives. (Scheman, 2001, p. 221)

The requirement of trustworthiness is not spread equally across all members of the society: greater demands and more onerous conditions are placed on policy makers, the experts of various kinds, the media, and medical carers, to name a few. This does not, however, let individual consumers of expertise off the hook. Consumers of information, in particular, have responsibility to show due diligence in accepting the testimony of their sources or in transmitting such information to others. In the remainder of the paper, we will briefly outline how elements of a climate of trust can address some of the worries listed here.

The idea of a climate of trust involves addressing the root causes of distrust not only in the specific domains where distrust manifests itself, but also the different power structures in which these domains are articulated. Establishing a climate of trust requires acknowledging the full range of the causes of distrust, which observing the social positions of the sceptics reveals to be more than merely intellectual; and this, in turn, means being prepared to change deeply rooted social and political structures and practices. If distrust is at least partly based on power inequalities and (past and current) discrimination, then consistent and visible efforts to rectify such inequalities are part and parcel of building a climate of trust.

Trust and distrust are valuable when exercised under the right conditions and directed towards the right individuals and organisations. The connecting element between them is in the vulnerability and disempowerment that are inherent in trust and are also at the root of the distrust arising from experiences of injustice. To trust, as we saw, is to make ourselves vulnerable but an added level of social vulnerability makes the act of trusting much more risky than it could otherwise be. We cannot remove the experience of vulnerability inherent the act of trusting , but in building a climate of trust we can work towards addressing the excesses of such feelings of vulnerability occasioned by experiences of social injustice.

Bibliography

Alfano, M. (2016). The topology of communities of trust. *Russian Sociological Review,* *15*(4), 30–56.

Almassi, B. (2012). Climate change, epistemic trust, and expert trustworthiness. *Ethics and the Environment 17*(2), 29–49. https://doi.org/10.2979/ethicsenviro.17.2.29.

Anderson, E. (2011). Democracy, public policy, and lay assessments of scientific testimony. *Episteme*, *8*(2), 144-164.

Arendt, H. (1972). *The crises of the republic: Lying in politics; Civil disobedience; On violence; Thoughts on politics and revolution*. Harcourt Brace & Co.

Baghramian, M., & Martini, C. (2018). Expertise and expert knowledge. [Special issue].*Social Epistemology*, *32*(6).

Baghramian, M., & Caprioglio Panizza, S. (2022). Scepticism and the value of distrust. *Inquiry*, 1–28. <https://doi.org/10.1080/0020174X.2022.2135821>

Baghramian, M., & Croce, M. (2021). Experts, public policy and the question of trust. In M. Hannon & J. De Ridder (Eds.). *Routledge handbook of political epistemology*. Routledge.

Baier, A. (1986). Trust and antitrust. *Ethics* *96*(2): 231–60. https://doi.org/10.1086/292745.

Baier, A. (2004). Demoralization, trust, and the virtues. In C. Calhoun (Ed.). *Setting the moral compass: Essays by women philosophers* pp. 176–189. Oxford University Press,.

Becker, L. C. (1996). Trust as noncognitive security about motives. *Ethics* *107*(1), 43–61.

Booth, K., & Wheeler N. (2007). *The security dilemma: Fear, cooperation and trust in world politics.* Palgrave Macmillan.

Brennan, J. (2020). Can novices trust themselves to choose trustworthy experts? Reasons for (reserved) optimism. *Social Epistemology*, *34*(3), 227-240.

Canovan, M. (1999). Trust the people! Populism and the two faces of democracy. *Political Studies,* *47* 2–16.

Code, L. (2006). *Ecological thinking: The politics of epistemic location*. [Studies in Feminist Philosophy]. Oxford University Press.

Cohen, M. A., & Dienhart J. (2013). Moral and amoral conceptions of trust, with an application in organizational ethics. *Journal of Business Ethics*, *112*(1): 1–13. doi:10.1007/s10551-012-1218-5.

Croce, M., & Baghramian, M. (2024). Experts – part I: What they are and how to identify them. *Philosophy Compass*, e13009. <https://doi.org/10.1111/phc3.13009>

Croce, M., & Baghramian, M. (2024). Experts – Part II: The sources of epistemic authority. *Philosophy Compass*, e70005. <https://doi.org/10.1111/phc3.70005>

Darwall, S. (2017). Trust as a second-personal attitude of the hearth. In P. Faulkner & T. Simpson (Eds.). *The philosophy of trust* (pp. 35-50). Oxford University Press.

Dellsén, F. (2018). The epistemic value of expert autonomy. *Philosophy and Phenomenological Research* (2), 344-361.

Desmond, H. (2022). Status distrust of scientific experts. *Social Epistemology,* *36* (5), 586-600.

Dewey, J. (1927). *The public and its problems*. Swallow Press.

Domenicucci, J., & Holton, R. (2017). Trust as a two-place relation. In P. Faulkner & T. Simpson (Eds.). *The philosophy of trust* (pp. 149-160). Oxford University Press.

Dormandy, K. (2020). Introduction: An overview of trust and some key epistemological applications. In Dormandy, K. (Ed.). *Trust in epistemology* (pp. 1-40). Routledge.

Douglas, H. (2013). The value of cognitive values. *Philosophy of Science*, *80*(5), 796–806. <https://doi.org/10.1086/673716>.

Facciolà, A., Visalli, G., Orlando, A., Bertuccio, M. P., Spataro, P., Squeri, R., Picerno, I., & Di Pietro, A. (2019). Vaccine hesitancy: An overview on parents' opinions about vaccination and possible reasons of vaccine refusal. *Journal of Public Health Research*, *8*(1), 1436. <https://doi.org/10.4081/jphr.2019.1436>

Farrell, D. M., & Suiter, J. (2019). *Reimagining democracy : Lessons in deliberative democracy from the Irish front line*. Cornell University Press. https://doi.org/10.7591/9781501749346)

Faulkner, P. (2007). On telling and trusting. *Mind*, *116*(464), 875-902.

Faulkner, P. (2014). The moral obligations of trust. *Philosophical Explorations*,*17*(3), 332-345.

Faulkner, P. (2015). The attitude of trust is basic. *Analysis* *75*(3), 424-429.

Faulkner, P. (2017). The problem of trust. In P. Faulkner & T. Simpson (Eds.). *The philosophy of trust* (pp. 110-129). Oxford University Press.

Feygina, I., Jost J. T., & Goldsmith, R. E. (2010). System justification, the denial of global warming, and the possibility of “system-sanctioned change.” *Personality and Social Psychology Bulletin,* *36*(3), 326–38. https://doi.org/10.1177/0146167209351435.

Foucault, M. (2003), *Society must be defended: Lectures at the Collège de France, 1975-76*. (M. Bertani, A. Fontana, F. Ewald, & D. Macey (Eds.). Picador.

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

Frost-Arnold, K. (2013). Moral trust and scientific collaboration. *Studies in History and Philosophy of* Science Part A, *44*(3), 301-310.

Furman, K. (2020). Emotions and distrust in science. *International Journal of Philosophical Studies* *28*(5), 713–30. <https://doi.org/10.1080/09672559.2020.1846281>.

Goldberg, Sanford C. (2020). Epistemically engineered environments. *Synthese,* *197*(7), 2783-2802.

Goldenberg, M. J. (2016). Public misunderstanding of science? Reframing the problem of vaccine hesitancy. *Perspectives on Science*, *24*(5), 552–81. https://doi.org/10.1162/POSC\_a\_00223.

Govier, T. (1992). Trust, distrust, and feminist theory. *Hypatia*, *7*(1), 16–33. doi:10.1111/j.1527-2001.1992.tb00695.x

Graham, P. J. (2012). Testimony, trust, and social norms. *Abstracta 6*(3), 92–116.

Grasswick, H. (2020). Reconciling epistemic trust and responsibility. In K. Dormandy (Ed.), *Trust in epistemology* (pp. 161-188)*.* Routledge.

Grasswick, H. (2014). Climate change science and responsible trust: A situated approach. *Hypatia 29*(3), 541–57. https://doi.org/10.1111/hypa.12090

Habermas, J.. (1985). *The theory of communicative action, Volume 1: Reason and the rationalization of society* (T. McCarthy, Trans.).

Hall, D. (2019). Internal reasons and the problem of climate change. *Theoria*, *66*(160). https://doi.org/10.3167/th.2019.6616003..

Hardin, R. (1993). The street-level epistemology of trust. *Politics and Society*, *21*(4), 505-529.

Hardin, R. (2002). *Trust and trustworthiness*. Russell Sage Foundation.

Harding, M. (2011). Responding to trust: Responding to trust. *Ratio Juris*, *24*(1), 75–87.

Hardwig, J. (1985). Epistemic dependence. *The Journal of Philosophy*, *82*(7), pp. 335-349.

Hardwig, J. (1991). The role of trust in knowledge. *The Journal of Philosophy*, *88*(12), 693-708.

Hawley, K. (2017). Trustworthy groups and organisations. In P. Faulkner & T. Simpson (Eds.). *The philosophy of trust*. Oxford University Press.

Hawley, K. (2019). *How to be trustworthy*. Oxford University Press.

Hempel, C. G. (1965). Science and Human Values. In *Aspects of scientific explanation and other essays in the philosophy of science*. (pp. 81–96). The Free Press.

Hollis, M. (1998). *Trust within reason*. Cambridge University Press. doi:10.1017/CBO9780511612244

Holton, R.. (1994). Deciding to trust, coming to believe. *Australasian Journal of Philosophy* *72*(1), 63–76. https://doi.org/10.1080/00048409412345881

Jasanoff, S. (2014). A mirror for science. *Public Understanding of Science*, 23 (1), 21–26. https://doi.org/10.1177/0963662513505509

John, S. (2018). Epistemic trust and the ethics of science communication: Against transparency, openness, sincerity and honesty. *Social Epistemology*, *32*(2), 75-87.

Jones, K. (1996). Trust as an Affective Attitude. *Ethics*, *107*(1), 4-25.

Jones, K. (2004). Trust and Terror. In p. DesAutels & M. Urban Walker (Eds.). *Moral psychology: Feminist ethics and social theory* (pp. 3-18)*.* Rowman and Littlefield.

Jones, K. (2013). Distrusting the trustworthy. In D. Archard, M. Deveaux, N. Manson, & D. Weinstock (Eds.). *Reading Onora O'Neill*. Routledge.

Kahan, D., Jenkins-Smith, H., & Braman, D. (2010). Cultural cognition of scientific consensus. *Journal of Risk Research*, *14*(2), 147-174.

Kelsall, R. (2021). The trust‐based communicative obligations of expert authorities. *Journal of Applied Philosophy* *38*(2), 288-305.

Kelsall, J., & Sorell, T. (2024), Two kinds of vaccine hesitancy. *Social Epistemology,* *39*(1), 40-55.

Koskinen, I. (2020). Defending a risk account of scientific objectivity. *British Journal for the Philosophy of Science* *71*(4), 1187–1207. https://doi.org/10.1093/bjps/axy053.

Kitcher, P. (2011). *Science in a democratic society*. Prometheus.

Krishnamurthy, M. (2015). Tyranny and the democratic value of distrust. *The Monist,* *98*(4), 391–406. https://doi.org/10.1093/monist/onv020

Kruger, J., & Dunning, D. (1999) Unskilled and unaware of it: How difficulties in recognizing one’s own incompetence lead to inflated self-sssessments. *Journal of Personality and Social Psychology*, *77*(6), 1121-1134.

Kuhn, T. S. (1977). Objectivity, value judgment, and theory choice. In *The essential tension: Selected studies in scientific tradition and change* (pp. 320–39). University of Chicago Press.

Lackey, J. (2018). Experts and peer disagreement.I In M. Benton, J. Hawthorne, & D. Rabinowitz (Eds.). *Knowledge, belief, and God: New insights in religious epistemology* (pp. 228-45). Oxford University Press.

Levy, N. (2019). Due deference to denialism: Explaining ordinary people’s rejection of established scientific findings. *Synthese*, *196*(1), 313-327.

Lewenstein, B., & Brossard, D. (2006). Assessing models of public understanding in ELSI outreach materials.. <https://doi.org/10.2172/876753>

Lombrozo, T., Thanukos, A., & Weisberg M. (2008). The importance of understanding the nature of science for accepting evolution. *Evolution: Education and Outreach,* 1, 290-298. 10.1007/s12052-008-0061-8

McGeer, V. (2008). Trust, hope and empowerment. *Australasian Journal of Philosophy*, *86*(2), 237–254. doi:10.1080/00048400801886413

McGeer, V., & Pettit, P. (2017). The empowering theory of trust (pp. 15–35). In P. Faulkner & T. Simpson (Eds.). doi:10.1093/acprof:oso/9780198732549.003.0002

McLeod, C. (2002). *Self-trust and reproductive autonomy.* MIT Press.

Miller, J. D. (2004). Public understanding of, and attitudes toward, scientific research: What we know and what we need to know. *Public Understanding of Science*, *13*(3), 273–94. https://doi.org/10.1177/0963662504044908

Miller, J. D. (1983). Scientific literacy: A conceptual and empirical review. *Daedalus*, *112* (2), 29–48.

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

Nguyen, C. T. (2018). Expertise and the fragmentation of intellectual autonomy. *Philosophical Inquiries,* *6*(2): 107–24.

Nguyen, C. T. (2022). Trust as an unquestioning attitude. Oxford Studies in Epistemology *7*, 214-244.

Nichols, T. (2017). *The death of expertise: The campaign against established knowledge and why it matters.* Oxford University Press.

Nickel, P. J. (2007). Trust and obligation-ascription. *Ethical Theory and Moral Practice*, *10*(3), 309-319.

Nickel, P. J. and Frank, L. (2020). Trust in medicine. In Judith Simon (Ed.). [*The Routledge handbook of trust and philosophy*](https://philpapers.org/rec/SIMRHO). Routledge.

Noonan, H. (2021). No trust is hybrid: reply to Faulkner. *Philosophia* Alfano.

O’Neill, O. (2002). *Autonomy and trust in bioethics*. Cambridge University Press. doi:10.1017/CBO9780511606250

Oberauer, K., & Lewandowsky S. (2016). Control of information in working memory: Encoding and removal of distractors in the complex-span paradigm. *Cognition,* *156*, 106–28. https://doi.org/10.1016/j.cognition.2016.08.007

Oreskes, N. (2019). *Why trust science*Princeton University Press.

Origgi, G. (2019). *Reputation: What it is and why it matters*. Princeton University Press.

Oughton, D. et al. (2004). An ethical dimension to sustainable restoration and long-term management of contaminated areas. *Journal of Environmental Radioactivity,* *74*(1-3), 171-183.

Popa, E. (2024), Values in public health: an argument from trust. *Synthese,* 203, 200. https://doi.org/10.1007/s11229-024-04650-8

Potter, N. N. (2002). *How can I be trusted? A virtue theory of trustworthiness*. Rowman & Littlefield.

Price, D. K. (1965). *The scientific estate.* Belknap Press of Harvard University Press.

Quine, W. V. (1970). On the reasons for indeterminacy of translation. *The* *Journal of Philosophy,* *67*(6), 178–83. https://doi.org/10.2307/2023887

Raz, J. (1986). *The morality of freedom*. Oxford University Press.

Raz, J. (1998). Disagreement in politics. *American Journal of Jurisprudence,* *43*(1), 25–52. https://doi.org/10.1093/ajj/43.1.25

Saunders, M. (2010). *Organizational trust: A cultural perspective*. Cambridge University Press.

Scheman, N. (2011). *Shifting ground: Knowledge and reality, transgression and trustworthiness.* Oxford University Press.

Simion, M. (2020). Testimonial contractarianism: A knowledge-first social epistemology. *Noûs*. https://doi.org/10.1111/nous.12337

Simpson, T. W. (2012). What is trust? *Pacific Philosophical Quarterly*, *93*(4), 550–569. doi:10.1111/j.1468-0114.2012.01438.x

Stanford, P. K. (2001). Refusing the devil’s bargain: What kind of underdetermination should we take seriously? *Philosophy of Science,* *68* (S3), 1–12. https://doi.org/10.1086/392893

Steinberg, J. (2017). Re-examining the early years of anti-retroviral treatment in South Africa: A taste for medicine. *African Affairs,* *116*(462), 60–79.

Sunstein, C. (2017). *#Republic: Divided democracy in the age of social media.* Princeton University Press.

Thomas, S. B., & Quinn, S. C. (1991). The Tuskegee Syphilis Study, 1932 to 1972: Implications for HIV education and AIDS risk education programs in the Black community. *American Journal of Public Health*, *81*(11), 1498–1505.

Turner, S. (2001). What is the problem with experts?’ *Social Studies of Science,* *31*(1): 123–49.

Walker, M. U. (2006). *Moral Repair: Reconstructing moral relations after wrongdoing*. Cambridge University Press.

Warren, M. E. (Ed.). (1999). *Democracy and Trust.*Cambridge University Press. doi:10.1017/CBO9780511659959

Webb, M. O. (1993). The epistemology of trust and the politics of suspicion. *Pacific Philosophical Quarterly,* *73*(4), 390–400. doi:10.1111/j.1468-0114.1992.tb00346.x

Wilholt, T. (2013). Epistemic trust in science. *British Journal for Philosophy of Science*, 64, 233-253.

Williams, B. (1979). Internal and external reasons. In R. Harrison (Ed.), *Rational action*, 101–13. Cambridge University Press.

Wilsdon, J., & Willis, R. (2004). See-through science: Why public engagement needs to move upstream. *DEMOS*.

Wylie, A. (2003). Why standpoint matters. In R. Figueroa & S. G. Harding (Eds.), *Science and other cultures: Issues in philosophies of science and technology* (pp. 26-48). Routledge.

1. This literature review is based on the White Paper on the ethics of trust in and the trustworthiness of experts, a deliverable of the Horizon 2020 Research and Innovation Project, PERITIA. Companion papers were published by Croce, M., & Baghramian, M. (2024). Work on the original White Paper was made possible by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 870883. More recently, additional work on this article by Baghramian was made possible by Horizon Europe ERA Chair Project ETICA under grant agreement ID:101187166. The views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union, who cannot be held responsible for them. [↑](#footnote-ref-1)
2. The possibility of betrayal is also one of the main reasons why trust is distinguished from mere reliance, where one cannot be said to be betrayed, but rather disappointed (see Baier, 1986; 1991). [↑](#footnote-ref-2)
3. However, sometimes what we rely on in trust is something other than goodwill, which has led to a number of objections to Baier’s goodwill-based account of trust, such as that goodwill is not necessary nor sufficient for trust (e.g., O’Neill (2002); Jones (2004)). [↑](#footnote-ref-3)
4. On cognitive vs. noncognitive trust, see Becker’s definition (1996, p. 44): “Let us call our trust ‘cognitive’ if it is fundamentally a matter of our beliefs or expectations about others' trustworthiness; it is noncognitive if it is fundamentally a matter of our having trustful attitudes, affects, emotions, or motivational structures that are not focused on specific people, institutions, or groups.” [↑](#footnote-ref-4)
5. Nguyen’s account is part of an original defence of trust that extends it to objects, including technology, removing the traditional requirement that the trustee be an agent. [↑](#footnote-ref-5)
6. On two-place Vs three-place trust, see Horsburgh (1960). [↑](#footnote-ref-6)
7. <https://rebellion.global/about-us/> [↑](#footnote-ref-7)
8. Domenicucci and Holton (2017, p. 151) [↑](#footnote-ref-8)
9. See also the discussion on trust in States and individuals representing States and other groups (e.g. Booth and Wheeler, 2008) and “organizational trust” (e.g., Saunders, 2010). [↑](#footnote-ref-9)
10. On the four pillars of medical ethics, see Beauchamp and Childress (2012). [↑](#footnote-ref-10)
11. The danger of over-reliance on experts was known to John Dewey who worried, already in the early 20th century, that the ever-increasing reliance on expert advice can diminish the scope of participation by ordinary citizens in the political process (Dewey, 1927). A worry that was echoed, in even stronger terms, by Hannah Arendt (1972) and other members of the Frankfurt School (Habermas, 1985) as well as Michel Foucault (2003) who, in their various ways, saw over-reliance on experts as a side-effect of the type of “scientism” that holds Western societies in its grip, and constitutes an inherently ideological stance masquerading under the banner of objectivity. See also Turner (2001). [↑](#footnote-ref-11)
12. See also the article “Against Epistocracy” by Åsa Wikforss on Epistocracy in this volume. [↑](#footnote-ref-12)
13. The research project of the Center for Ethics and Public Affairs (ETICA) will rely on this methodology for assessing and fostering trust in a time of crisis. [↑](#footnote-ref-13)
14. We should point out, maybe unnecessarily, that the mere uncertainty of the sort inherent in the methodology and practice of science, and readily acknowledge by scientists, is not a sufficient reason for scpeticism about scientific expert advice. [↑](#footnote-ref-14)
15. At the same time, it is important to take values into account in a way that does not deny their incompatibility with others and that does not distort the scientific message. Some values are just incompatible with pro-environmental behaviour and a just system. That is why, the social situatedness of the untrusting needs to be taken into account, not only by experts who seek trust, but also by the untrusting themselves. This is Heidi Grasswick's (2014) proposal to address climate scepticism. [↑](#footnote-ref-15)
16. See also, Baghramian and Panizza (2022) for a discussion of this topic. [↑](#footnote-ref-16)
17. The approach is tantamount to taking a ‘situated’ perspective, advocated by feminist epistemologists such as Code (2006) and Wylie (2003). [↑](#footnote-ref-17)