20

TRUST IN INSTITUTIONS AND GOVERNANCE

Mark Alfano and and Nicole Huijts

20.1 Introduction

Elaborating on themes from Hobbes (1668/1994), Alfano (2016a) has argued that warranted trust fosters multiple practical, epistemic, cultural and mental health goods. In this paper, we focus on the practical and epistemic benefits made possible or more likely by warranted trust. In addition, we bear in mind throughout that trusting makes the trustor vulnerable: the trustee may prove to be unlucky, incompetent or an outright betrayer. With this in mind, we also focus on warranted lack of trust and outright distrust, the benefits they make possible, and the harms that the untrusting agent is protected against and may protect others against (see also D'Cruz, this volume). We use cases of (dis)trust in technology corporations and the public institutions that monitor and govern them as examples throughout this chapter.

In so doing, we build on the accounts of Jones (2012) and Alfano (2016a, 2016b) to develop definitions of various dispositions related to trusting and being trusted. Our goal is then to argue that private corporations and public institutions have compelling reasons both to appear trustworthy and to actually be trustworthy. From this it follows that corporations and institutions have strong instrumental and moral reasons to adopt a suite of policies that promote their appearing trustworthy and being trustworthy.

Here is the plan for this chapter: first, we explain the conception of trustworthiness that we employ. We model trustworthiness as a relation among a trustor, a trustee and a field of trust defined and delimited by its scope. In addition, both potential trustors and potential trustees are modeled as being more or less reliable in signaling either their willingness to trust or their willingness to prove trustworthy in various fields in relation to various other agents. Second, following Alfano (2016a) we argue that the social scale of a potential trust relationship partly constrains both explanatory and normative aspects of the relation. Most of the philosophical literature focuses on dyadic trust between a pair of agents (Baier 1986, Jones 2012, McGeer 2008, Pettit 1995), but there are also small communities of trust (Alfano 2016a) as well as trust in large institutions (Potter 2002; Govier 1993; Townley and Garfield 2013, Hardin 2002). The mechanisms that induce people to (reasonably) extend their trust vary depending on the size and structure of the community in question. Mechanisms that work in dyads and small communities are often unavailable in the context of trusting an

institution or branch of government. Establishing trust on this larger social scale therefore requires new or modified mechanisms. In the third section, we recommend several policies that tend to make institutions more trustworthy and to reliably signal that trustworthiness to the public; we also recommend some ways to be intelligently trusting (see also O'Neill, this volume). We conclude by discussing the warrant for distrust in institutions that do not adopt the sorts of policies we recommend; warranted distrust is especially pertinent for people who belong to groups that have historically faced (and in many cases still do face) oppression (see also Medina as well as Potter, this volume).

20.2 A Framework for Global, Rich Trust

To trust someone is to rely on them to treat your dependency on them as a compelling if not universally-overriding reason to act as expected. As Jones (2012) and Alfano (2016b) emphasize, trusting and being trusted are conceptually and developmentally interlocking concepts and phenomena (see also Scheman, this volume). Jones and Alfano also agree that trusting and being trusted are always relative to a domain. Alfano (2016a) glosses the domain as a field of valued practical concern and activity that is defined and delimited by its scope. To move from the descriptive phenomena of trusting and being trusted to the normative phenomena of being trustworthy and trusting (i.e., situations in which evaluations of appropriateness or warrant play a part), we need a theory of when trust is warranted. This would allow us to say, schematically, that B is trustworthy in domain D to the extent that she possesses a disposition that warrants trust, and that A is trusting in domain D to the extent that he possesses a disposition to extend trust when it is warranted.

We begin with Jones's (2012) partner-relative definition:

> B is trustworthy with respect to A in domain of interaction D, if and only if she is *competent* with respect to that domain, and she *would* take the fact that A is counting on her, *were* A to do so in this domain, to be a *compelling* reason for acting as counted on.

Next, we extend Jones's account of partner-relative *rich trustworthiness* by articulating congruent concepts of *global rich trustworthiness, partner-relative rich trustingness* and *global rich trustingness*. Jones points out that one agent's being trustworthy without anyone being able to tell that she is trustworthy is inefficient. Such a person may end up being trusted haphazardly, but her dependency-responsiveness will go almost largely unnoticed, unappreciated and unused.

This leads to a pair of unfortunate consequences. First, people who would benefit from depending on a trustworthy person in the relevant domain will be left at sea, forced either to guess whom to trust or to try to go it alone. Second, the trustworthy person will not receive the credit, esteem, resources and respect that come with being trusted. Things would go better for both potential trustors and trustworthy people if the latter could be systematically distinguished from the untrustworthy. This leads Jones (2012:74) to articulate a conception of partner-relative rich trustworthiness as follows:

> B is richly trustworthy with respect to A just in case (i) B is willing and able reliably to signal to A those domains in which B is competent and willing to take the fact that A is counting on her, were A to do so, to be a compelling

reason for acting as counted on and (ii) there are at least some domains in which she will be responsive to the fact of A's dependency.

Building on this definition, we can further define *global* (i.e. non-partner-relative) *rich trustworthiness* as follows:

> For all agents, B is globally richly trustworthy to the extent that (i) B is willing and able reliably to signal to others those domains in which B is competent and willing to take the fact that others are counting on her, were they to do so, to be a compelling reason for acting as counted on and (ii) there are some domains in which she will be responsive to others' dependency.

Global rich trustworthiness is a generalization of rich trustworthiness. It can be construed as the aggregate signature of rich trustworthiness that B embodies. Global rich trustworthiness measures not just how B is disposed towards some particular person but how B is disposed towards other people more broadly. It is defined using "to the extent that" rather than the biconditional because it is impossible for anyone to be globally richly trustworthy towards the entire universe of potential partners. Global rich trustworthiness therefore comes in degrees on three dimensions: partner (whom she is trustworthy towards), field (in what domains she is trustworthy), and extent (how compelling she finds the dependency of particular others in given fields).

Congruent to partner-relative rich trustworthiness, we can also define *partner-relative rich trustingness*:

> A is richly trusting with respect to B just in case (i) A is willing and able reliably to signal to B those domains in which A is willing to count on B to take A's counting on him to be a compelling reason for him to act as counted on and (ii) there are some domains in which A is willing to be dependent on B in this way.

Partner-relative rich trustingness is indexed to a particular agent. I might be richly trusting towards you but not towards your skeezy uncle. Just as it is important for potential trustors to be able reliably to identify trustworthy partners, so it is important for trustworthy partners to be able reliably to identify people who are willing to extend trust. Not only does this save the trustworthy time and effort but also it may empower them to accomplish things they could not accomplish without being trusted. For instance, an entrepreneur needs to find trusting investors to get her new venture off the ground, which she can do more effectively if trusting investors signal that they are willing to be dependent on the entrepreneur to use their money wisely and repay it on time and in full.

Aggregating rich trusting dispositions allows us to define *global rich trustingness*:

> For all agents, A is globally richly trusting to the extent that (i) A is willing and able reliably to signal to others those domains in which A is willing to count on them to take A's counting on them to be a compelling reason for acting as counted on and (ii) there are some domains in which A is willing to be dependent in this way.

Like global rich trustworthiness, global rich trustingness is a generalization of its partner-relativized cousin. It measures not just how A is disposed towards some particular person but how A is disposed towards other people more broadly. And like global rich trustworthiness, it is defined using "to the extent that" rather than the biconditional because it is impossible for anyone to be globally richly trusting towards the entire universe of potential partners. Global rich trustingness is therefore parameterized on the dimensions of partner (whom she is willing to be trusting towards), field (in what domains she is willing to be trusting), and extent (how compelling she expects her dependency to be for others).

20.3 The Social Scale of Trust

With these definitions in hand, we now turn to the social scale of trust. As Alfano (2016a) points out, restricting discussions of trust to the two extremes of the social scale (dyadic trust vs. trust in large, faceless institutions and governments) ignores all communities of intermediate size. As work on the psychological and neurological limits of direct sociality shows (Dunbar 1992), there are distinctive phenomena associated with trust at intermediate scales, which can sometimes be modulated to apply at the largest scales. Attending to the full spectrum enables us to see elements of continuity as well as breaks in continuity. Our discussion in this section is framed by the following questions:

(trustworthiness-1) How can people become responsive to the dependency of others?
(trustingness-1) How can people become willing to rely on the dependency-responsiveness of others?
(trustworthiness-2) How can people reliably signal their dependency-responsiveness?
(trustingness-2) How can people reliably signal their willingness to rely on the dependency-responsiveness of others?

Answering these questions points us in the direction of policies and practices that people and institutions can adopt to better approximate partner-relative and global rich trustworthiness and trustingness.

We will follow Alfano (2016a) by conceptualizing humans and their relations as a directed network, in which nodes represent agents and edges represent channels for both actions (e.g. communication, aid, harm) and attitudes (e.g. belief, knowledge, desire, aversion, trust, distrust). In this framework, X can become responsive to the dependency of Y only if there is one or more *short epistemic geodesic* (shortest path of communicative and epistemic edges) from X (through other nodes) to Y, giving X first-, second- … or nth-hand knowledge of Y's beliefs, desires, aversions, and so on. The longer the epistemic geodesic, the more opportunities for noise or bias to creep into the chain of transmission and thus the higher the likelihood that X will not reliably come to understand Y's epistemic and emotional perspective.

Even if X does reliably come to understand Y's epistemic and emotional perspective, this does not, on its own, guarantee that X will be responsive to Y's dependency. Recall that responsiveness is defined in this context in terms of treating someone's dependency as a compelling reason to act as counted on. X could know full well about Y's needs, preferences, dependencies and fears without treating these as a reason to act, let alone a compelling reason to act. This is similar to the insufficiency of empathy to motivate compassionate action (Bloom 2016): knowing is half the battle, but it is only half the battle. Motivation is needed as well.

One source of the needed motivation is *goodwill*, as Baier (1986) has pointed out. Goodwill is established and maintained through activities like social grooming (Dunbar 1993), laughing together (Dezecache and Dunbar 2012), singing and dancing together (Dunbar 2012), or enduring traumatic loss together (Elder and Clipp 1988). Jones (2012) persuasively argues, however, that there can be motivational sources other than goodwill. One important additional motivator is concern for *reputation*, which is especially pertinent when one is embedded in a social structure that makes it likely that others will achieve mutual knowledge of how one has acted and for what reasons one has acted (Dunbar 2005; see also Origgi, this volume). One-off defection or betrayal in an interaction exposes one to loss of reputation and thereby to exclusion from the benefits of further cooperation and coordination (see also Dimock, this volume). In a community with short epistemic geodesics, reputation-relevant information is likely to travel quickly. Dunbar (1993) estimates that at least 60% of human conversational time comprises gossip about relationships and personal experiences. As Alfano and Robinson (2017) argue, these phenomena make the disposition to gossip well (to the right people, about the right people, at the right time, for the right reason, etc.) a sort of virtue: a disposition that protects both oneself and other members of one's community from betrayal while punishing or ostracizing systematic defectors.

This brings us to an epistemic benefit of small-world communities (Milgram 1967), which are characterized by sparse interconnections but short geodesics (due to the presence of hubs within local sub-communities). Such communities are highly effective ways of disseminating knowledge. In the case of gossip and related forms of communication, the information in question concerns the actions, intentions and dispositions of another person. In computer science, it has been shown that, depending on the topology of a communicative network, almost everyone gets the message even when the probability of any particular agent gossiping is only between 0.6 and 0.8 (Haas et al. 2006).

There are two main reasons that such communities foster knowledge. First, because they effectively facilitate testimonial knowledge-transfer, they make it likely that any important information in the community eventually makes the rounds. Second, to the extent that the members of the community have at least an implicit understanding of the effectiveness of their own testimonial network, they are in a position to achieve second- and even higher-order levels of mutual knowledge. They can reasonably make inferences like, "If that were true, I would have heard it by now" (Goldberg 2010). They might also go further by making judgments like, "Because that's true, everyone in my community must have heard it by now."

In addition to goodwill and reputation, *solidarity* – understood here in terms of individuals sharing interests or needs and taking a second-order interest in each other's interests (Feinberg 1968), typically accompanied by self-identification with their group's accomplishments and failures – can motivate someone to respond to the dependency of another. Such self-identification with a group informs our self-conceptions. It gives us a sense of belonging, home and history (Nietzsche 1874 / 1997). It provides us with heroes and villains on whom to model our behavior and practice moral judgments. It helps to cement bonds within a community.

If this is on the right track, then a partial answer to **(trustworthiness-1)** is that people become responsive to dependency of others by being connected by short epistemic geodesics along with some combination of goodwill (fostered by in-person interaction), desire to maintain a good reputation (fostered by a small-world epistemic network), and solidarity between those in a dependent position and those with more power. The discussion so far also gives us a partial answer to **(trustingness-1)**. It makes sense to rely

on the dependency-responsiveness of another person to the extent that one is connected by a short epistemic geodesic, one has interacted positively with them in the past, one has received positive and reliable reputational information about them, and one expects them to feel a sense of solidarity.

What about **(trustworthiness-2)** and **(trustingness-2)**? These questions relate not to being trustworthy and trusting, but to being richly trustworthy and trusting. We contend that people come to reliably signal their dependency-responsiveness in two main ways. First, they can be transparent about their reasoning processes (not just the outcomes of these processes), which will showcase which reasons they are sensitive to in the first place (and where they have moral blindspots – see DesAutels 2004) and which among the reasons they are sensitive to they typically find compelling. Second, they can solicit other agents who are already trusted by third parties to vouch for them. Such vouching can lead third parties to extend their trust. In the first instance, they enable X to trust Y through some mediator M. More generally, transitive chains of trust may help X to trust Y through M_1, M_2 … M_n, and shorter chains can in general be expected to be more robust.

Small-scale groups in which everyone knows everyone can sustain the transitivity of trust among all their members. As the size of community increases, however, the need for vicarious or mediated trust increases. X vicariously trusts Y through M with respect to field of trust F just in case X trusts M with respect to F, M trusts Y with respect to F, and X trusts M's judgment about who is trustworthy with respect to F. Vicarious trust has a distinctive counterfactual signature in the sense that, if X vicariously trusts Y through M, then were X to become directly acquainted with Y, X would continue to trust Y non-vicariously. We can think of this in terms of *delegation* (empowering someone to extend your trust vicariously) and *ratification* (explicitly confirming an instance of delegation). In cases where acquaintance with Y leads X to withdraw rather than ratify her vicarious trust in Y, she may also begin to doubt M. To illustrate, suppose my boss trusts me to complete a task, and that I sub-contract out a part of that task to someone she distrusts. If she finds out that I have done this, she will most likely withdraw her trust from me – at least regarding this task and perhaps more generally.

Shy of such a highly demanding approach to transitivity, we might ask about extending one's trust one or two steps out into a community (Figure 20.1). What reasons are there for C to trust D, who is trusted by someone she trusts (B)?

In addition to *delegation*, we might focus on the phenomenon of *vouching*. B vouches for D to C if B makes himself accountable for any failure on D's part to prove trustworthy.
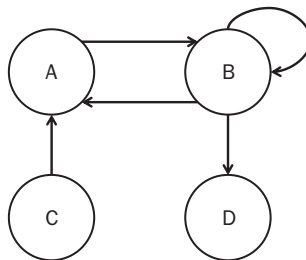

Figure 20.1 Trust Networks

How is such vouching meant to work? It relies on the rationality of extending trust transitively, at least in some cases. In other words, it relies on the idea that, at least sometimes, if C trusts B and B trusts D, then C has a reason to trust D. This reason need not be compelling. C can withhold her trust from D even as she gives it to B. Our hypothesis is that transitivity provides a *pro tanto* reason to extend trust, not an all-things-considered reason. There are two arguments for this hypothesis. First, competence in a domain is highly associated with meta-competence in making judgments about competence in that domain (Collins and Evans 2007: chapter 2). If C trusts B, that means C deems B competent with respect to their shared field of trust. It stands to reason, then, that C should expect B to be better than the average person at judging the competence of others in that field. So if B gives his trust to D, C has a reason to think that D is competent (see also Miller and Freiman, this volume).

Second, it is psychologically difficult and practically irrational to consciously engage in efforts to undermine your own values in the very process of pursuing and promoting those values. Imagine someone locking a door while they are trying to walk through the doorway. Someone could perhaps do this as a parapraxis. Someone could do it as a gag, or in pretense, but it is hard to envision a case in which someone does this consciously. Likewise, it is hard to envision a case in which someone is genuinely dependency-responsive, and consciously expresses that responsiveness by recommending that you put your fate in the hands of someone they expect to betray your trust. They might do so by mistake, as a gag or in pretense, but a straightforward case is difficult to imagine. If C trusts B, that means C judges that B is responsive to C's dependency. It stands to reason, then, that C should expect B to act on that responsiveness in a practically rational way. So if B gives his trust to D, C has a reason to think that D would act consistently with B's responsiveness to C's dependency.

Putting these together, if C trusts B and B trusts D (with respect to the same field of trust), then C has a reason to think that D is competent and responsive to the dependency of people like C. In other words, C has *pro tanto* reasons to trust D.

On the question of rich trustingness, we see two main ways to signal it. First, the agent could establish a track-record of trusting particular types of people in particular domains to particular extents. This track-record would then be the basis of a reputation for having a signature of trusting dispositions (relativized to partners, domains and extent of trust). This leaves open, however, how the first step is to be taken. How can people with no reputation – good or bad – go about establishing their rich trustingness? This brings us to our second method of signaling. Someone can begin to establish a record of trustingness by engaging in small "test" dependencies: extending her trust just a little bit even when she lacks compelling reasons to do so. Such tests simultaneously enable the trustor to establish her reputation and provide her with feedback about the trustworthiness of others. Doing so might seem reckless, but if it is viewed from the point of view of information-gathering (I trust you in order to find out what kind of person you are rather than to reap the direct benefits of trust) this strategy is sensible for people who have enough resources to take small risks.

20.4 Rich Trustworthiness for Institutions: Policy Recommendations

Some of the mechanisms that (reasonably) induce people to extend their trust vary depending on the size of the community envisioned; the ways in which (rich) trustworthiness and trustingness can be established vary with these mechanisms. Shaming, shunning, laughing together, building and maintaining a good reputation over the long

run through networks of gossip, bonding over shared enjoyment and suffering, sharing a communal identity: such phenomena can warrant trust in a dyad or small community. However, with a few exceptions such as reputation-management, they are typically unavailable in the context of trusting an institution or branch of government. In addition, it may be possible for such institutions to reliably signal their trustworthiness only to some stakeholders – namely, those with whom they have already established a good-enough reputation as opposed to those who have experienced a history of neglect and betrayal. Establishing global rich trustworthiness in institutions and governance may require new mechanisms or policies.

Reputation is one mechanism that applies at both the individual/small-group level and the institutional level. As institutions have a long lifespan, it is possible that, in new trust-requiring situations, past actions or omissions have seriously damaged trust. This situation is not easy to repair, and reliable signaling might not be possible in the face of distrust. Furthermore, the public – and especially groups that have suffered from oppression or discrimination – can be reasonable in distrusting actors that have proven untrustworthy in the past (Krishnamurthy 2015). The case of opposition of the Standing Rock Sioux to the Dakota Access Pipeline (Plumer 2016) is a good example, as are various nuclear siting controversies in the United States (Kyne and Bolin 2016; Wigley and Shrader-Frechette 1996). We thus do not advocate maximizing trust, but fine-tuning it.

When a distrusted institution has, however, changed its ways and has become trustworthy, it can help to include other more trusted actors in decision-making or information provision to signal this trustworthiness. More trusted partners could potentially vouch for less trusted partners which might then lead to higher trust in the whole consortium. The fact that often parties risk their own reputations by vouching for a former offender means that they have an incentive to do so only when they have very good reason. Huijts et al. (2007) showed with a survey study about a hypothetical carbon capture and storage (CCS) project that trust in involved actors together to make the right decisions and to store CO_2 safely and responsibly was predicted by trust in three different actors: government, industry and environmental NGOs. Trust in government was rated higher than trust in the industry, and had a much larger effect on overall trust than trust in industry. This suggests that in some cases considerable involvement of and transparent oversight by the government (assuming, of course, that the government itself is trusted) can help to overcome low levels of trust in other institutions.

In this section, we describe mechanisms that can build or undermine trust (and the reliable signaling thereof) in the context of institutions and governance, using large, potentially risky, energy technology projects as an exemplary case. Energy technology projects are regularly proposed to increase energy security or reduce environmental problems such as air pollution and climate change. Examples are windmill parks, CCS, high-voltage power lines, shale gas mining, geothermal energy and nuclear power plants. Although these energy projects can offer important benefits for society and the environment, they also introduce potential risks and drawbacks (e.g., visual intrusion, increased traffic and risks of oil spills and nuclear meltdowns). Trust in institutions involved with the implementation of energy technologies, such as energy companies and governmental regulatory bodies, is an important predictor for citizens' evaluation of and responses to implementations of such technologies (Huijts et al. 2012; L'Orange Seigo et al. 2014). Higher trust in those responsible for the technology is generally associated with higher perceived benefits, lower perceived risks and more

positive overall evaluations of a technology. By contrast, when people do not trust governmental institutions and private companies to manage these risks and drawbacks responsibly, the projects are likely to be contested (Huijts et al. 2014), as we have seen recently in the opposition of the Standing Rock Sioux to the Dakota Access Pipeline (Plumer 2016).

Institutions and the governmental agencies that oversee them thus have a need not only to be trustworthy but to reliably signal to affected populations that they are trustworthy. They need to approximate as closely as possible global rich trustworthiness. However, they cannot easily rely on the processes that build trust in dyads and small communities. Governmental regulators and representatives of the industries they regulate cannot be expected to bond through laughing and crying together with all stakeholders. Personally identified ambassadors can give a face to the institutions they represent, but this is typically more a matter of marketing than reliable signaling. Empirical research suggests that trust in parties responsible for a technology is based on their perceived competences and intentions (Huijts et al. 2007). Knowing this, companies could insist that they have good intentions and competence, but such a direct approach is liable to fail or backfire. "Would I lie to you?" is not an effective retort to someone who has just questioned whether you are lying to them. Indeed, Terwel et al. (2009) showed that trust in an institution is lower when it provides communication about CCS that is incongruent with the motives it is presumed to have (e.g. involving environmental arguments for companies and economic arguments for environmental NGOs) than when it gives arguments that are congruent with inferred motives. Perceived honesty was found to explain the effect of perceived congruence on trust. Directly insisting on one's own good intentions when one is not already perceived as honest is thus not suitable for building and gaining trust.

We now turn to three more indirect ways for public and private institutions to reliably signal trustworthiness to a diverse group of stakeholders (i.e. to approximate rich global trustworthiness), which we label *voice, response* and *diversity*. The first way is to give stakeholders a meaningful voice in the decision-making process. Allowing voice by different parties is epistemically valuable. As standpoint epistemologists have argued, different parties bring different values and background knowledge, and victims of oppression often have distinct epistemic advantages when it comes to the vulnerabilities of those in dependent situations (Harding 1986; Wylie 2003). Since responsiveness to dependency and vulnerability is an essential component of trustworthiness, including such people in the decision process is likely to improve decision-making.

Empirical research has shown that allowing voice can indeed increase trust. Terwel et al. (2010; study 3) showed that respondents reported more trust in the decision-makers and more willingness to accept the decision about implementing a CCS project when the public was also allowed a voice in the decision-making procedure as compared to when only the industry and environmental NGOs were allowed a voice. Not only voice by citizens but also voice by other parties can affect trust and acceptance. Study 1 in the same paper showed that allowing voice by industry and environmental NGOs led to higher trust in the decision-makers and higher acceptance of the decision outcome than when no voice was allowed. Furthermore, when only one other party was allowed to voice a view, even when this was a trusted one (e.g. environmental NGOs), this did not lead to higher trust. Only allowing voice to all of the dissimilar parties (industry and environmental NGOs) was found to lead to higher trust in the decision-makers

and higher acceptance of the decision outcome (study 2). The authors argue that allowing different parties to voice their opinion speaks of procedural fairness, which increases trust.

Of course, there is a danger that allowing voice is done in a purely instrumental way, as "window dressing." However, if that becomes apparent it could substantially undermine trust by signaling that those with the power to make decisions are not actually responsive to the dependency of those who have been asked to trust them. There are thus pragmatic as well as ethical and epistemic reasons to genuinely involve stakeholders in decision-making. This provides for short epistemic geodesics and hence knowledge of people's particular dependencies, and means that decision-makers genuinely engage in an exchange between equal peers (Habermas 1990) rather than broadcasting a message without engaging in substantive dialogue.

In order to be able to judge a socio-technical proposal and to take part in decision-making, citizens need to be able to gather relevant information. Open, timely and respectfully provided information is important for how citizens perceive a decision-making process. Industry managers and policy makers should not only communicate what they think citizens should know about the technology but also provide information that citizens are particularly concerned about and interested in. Short geodesics are needed to create awareness of what citizens are concerned about and interested in and to design matching information. Response – the provision of information that is open, honest, timely, respectfully provided, and suited to the concerns and interests of the public is therefore the second way for public and private institutions to signal trustworthiness. In so doing, they establish not only their trustworthiness but also their rich and (to the extent that they signal successfully to all affected stakeholders) global trustworthiness.

The third way for public and private institutions to signal trustworthiness, thus establishing their rich global trustworthiness, is by creating diversity within the institution at all levels but especially the top levels where the most important decisions are made. Empirical studies suggest that higher perceived similarity in goals and values between oneself and those responsible for a technology leads to higher levels of trust. For example, Siegrist et al. (2000) showed for nuclear power that a higher perceived similarity between oneself and company managers with respect to values, goals, behaviors, thoughts and opinions goes together with higher trust in those responsible for the technology. Huijts et al. (2007) similarly showed that higher perceived similarity in important goals and in the way of thinking between representatives of the actor and oneself correlates with higher trust in actors involved with a new technology (CCS in this case). Having more diversity in institutions is likely to create a situation in which citizens and other stakeholders can point to an empowered individual who is relevantly similar to them, thereby fostering a sense of solidarity.

The effect of perceived similarity is not just a bias. As feminist epistemologists argue, people who are similar to you are likely to have better epistemic familiarity with your problems, concerns and values (Daukas 2006; Grasswick 2017) and are also likely to share more of your values. Embracing diversity can thus serve as a source of epistemic improvement.

In addition, the involvement of diverse parties in information provision can be helpful in creating trust in another way. When different institutions independently provide information, this may increase the likelihood that each citizen trusts at least one source of information, which can help them form an opinion. However, this can

easily lead to polarization. When institutions formulate common information texts, this may be even more helpful, as then one piece of information is available that is checked and approved by parties with different value-bases and different interests. Indeed, such a process approximates the best-known way to harness the wisdom of crowds by aggregating the information and values of independent, diverse sources (Surowiecki 2005; see also Ter Mors et al. 2010).

The benefits of the diversity of decision-makers can only be reliably signaled, however, when citizens and other stakeholders are aware of this diversity and see at least some of the decision-makers as standing in solidarity with them. To increase awareness of diversity, it is necessary to make it visible. Meijnders et al. (2009) showed that trust in information provided by a journalist about genetically modified apples became higher when the journalist was expressing an attitude about something that was congruent with the attitude of the respondent, independent of whether this congruent attitude was about a similar technology (genetically modified oranges), or about a different technology (a cash machine with voice control). Judgments of similarity between oneself and the journalist were found to mediate this effect. This shows that an awareness of some kind of similarity (of a particular opinion in this case) can indeed increase trust.

Of course, these are not the only ways that institutions can signal trustworthiness. We are not offering an exhaustive list in this short chapter, but we believe that the policies suggested here do not enjoy sufficient appreciation. Attending to the phenomena we have canvassed in this section is part of what an agent needs in order to extend their trust intelligently. Moreover, by demonstrating their appreciating of *voice, response* and *diversity*, potential trustors can signal their willingness to trust (in certain conditions) and thus become richly trusting.

20.5 Trust and Distrust in Other Institutions

Thus far, we have focused on trust and distrust in corporations and governmental agencies that regulate and oversee them. Naturally, there are other relevant institutions when it comes to warranted trust and distrust, such as the military, universities, hospitals, courts and churches and other religious institutions. As we lack the space to address all the differences among these institutions in this chapter, we here make just a few remarks about them.

First, these other institutions face the same challenges in establishing trust that corporations and government agencies do. They are too large to rely on practices like social grooming, laughing together, crying together, and so on. Second, like the institutions we have focused on, they can (and should) rely on reputation-building and reputation-management mechanisms. Third, the mechanisms of *voice, response* and *diversity* should in general work just as well for these institutions as they do for corporations and governmental regulators. There may, however, be some exceptions. For example, it may not be appropriate to give equal voice in decision-making about policing to criminals and law-abiding citizens. In addition, sometimes fully embodying the ideal of *response* is impossible because doing so would violate privacy rights or other rights (e.g. in a hierarchical command structure such as the military). Finally, if our assumption of parity is on the right track, then lack of *voice, response* and *diversity* indicates that it is reasonable for many stakeholders to respond to many contemporary institutions (e.g. the Catholic Church and its all-male priesthood, along with racially segregated police forces) with distrust rather than trust. This leads us to our final section.

20.6 Warranted Distrust

While this chapter has focused primarily on the ways in which warranted trust can be established and trustworthiness reliably signaled, we want to end with a note of warning: citizens and other stakeholders (especially those who have suffered neglect or betrayal by institutions) may become reasonably distrusting of government and industry in several ways. Recognizing this fact and the difficulty of reliably signaling trustworthiness to such potential trustors is an important and under-explored responsibility of institutions. We will discuss a number of examples.

First, people may become (reasonably) distrustful when authorities have a preset agenda and at most do public consultation as a form of window dressing, if at all. In the case of a wind farm in Australia, citizens reported that they felt they were not heard in the decision-making process, which sparked opposition to the project (Gross 2007). Some residents living near a newly built powerline in the Netherlands reported that they were given a false sense of influence. The interviewed citizens thought they were heard to avoid civil unrest, but that they did not have an actual influence on the decision-making (Porsius et al., 2016). Also in a CCS case in the Netherlands, local parties and citizens protested to influence a situation in which they were not given formal influence (Brunsting et al. 2011; Feenstra et al.2010:27). Brunsting et al. (2011:6382) concluded that, "The timing of public involvement reinforced the impression that Shell would be the only beneficiary which was therefore not a highly trusted source of information about safety or costs and benefits."

Second, people may become justifiably distrustful when authorities select their experts in a way that appears to be biased to support the view of the authority. In the Dutch CCS case, it was claimed that a critical report by a university professor had been held back from the decision-making process, which generated negative media attention and questions in parliament (Feenstra et al.2010:25).

Third, people may become distrustful when important values are ignored because there is no room for genuine exchange of emotions, values and concerns. In the Dutch CCS case, fear about genuine risks had been labeled emotional and irrational, meaning that such fears were silenced (Brunsting et al. 2011). This hampers respectful exchange of values and concerns. In the case of the wind farm project in Australia, interviewed citizens complained of greed and jealousy related to the fact that the person who owns the land on which a windmill is placed gains substantial income from it, while those living nearby suffer from drawbacks such as visual intrusion and noise annoyance but receive inadequate compensation (Gross 2007). If these fairness considerations had been taken into account earlier on in the project, the outcomes of the project would have likely been more acceptable.

Fourth, improper information provision can hamper opportunities to come to a genuine exchange of emotions, values and concerns and may lead to distrust in those responsible for the technology. In several cases, information was provided too late or was not targeted to the audience's interests. For example, in the Dutch CCS case, at the start of the project no information was available that was tailored to the public, that sought information about local costs and benefits, and that was endorsed by multiple parties (Brunsting et al. 2011). This happened only later in the project when it is was likely too late to make a difference. Also around the implementation of the high-voltage powerline in the Netherlands, citizens perceived a lack of transparency; they lacked personally relevant, timely and respectful information provision, which was associated with a lack of trust in the information

provision (Porsius et al. 2016). For the Australian wind farm project, a lack of clear notification of and information about the project at the start of the project was reported to spark opposition (Gross 2007).

Fifth, people may also become distrusting and take opposing actions when only arguments framed in technical language are allowed in the arena, thereby favoring experts at the expense of stakeholder involvement (cf. Cuppen et al. 2015) and when the boundaries of a debate are set in such a way that important alternatives are already excluded at the outset. Both these problems were reported to be the case in the heavily contested CCS case in the Netherlands (Cuppen et al. 2015; Brunsting et al. 2011).

So, while trust may often be a good thing, it needs to be earned. When corporations, institutions and governments do not make serious and public efforts both to be and to appear trustworthy, it is reasonable for citizens to react with distrust and take action to prevent the implementation of risky new technologies.

References

Alfano, M. (2016a) "The Topology of Communities of Trust," *Russian Sociological Review* 15(4): 30–56.

Alfano, M. (2016b) "Friendship and the Structure of Trust," in A. Masala and J. Webber (eds.), *From Personality to Virtue: Essays in the Psychology and Ethics of Character*, Oxford: Oxford University Press.

Alfano, M. and Robinson, B. (2017) "Gossip as a Burdened Virtue," *Ethical Theory and Moral Practice* 20: 473–482.

Baier, A. (1986) "Trust and Antitrust," *Ethics* 96: 231–260.

Bloom, P. (2016) *Against Empathy: The Case for Rational Compassion*, New York: HarperCollins.

Brunsting, S., De Best-Waldhober, M., Feenstra, C.F.J. and Mikunda, T. (2011) "Stakeholder Participation Practices and Onshore CCS: Lessons from the Dutch CCS Case Barendrecht," *Energy Procedia*, 4: 6376–6383.

Collins, H. and Evans, R. (2007) *Rethinking Expertise*, Chicago, IL: University of Chicago Press.

Cuppen, E., Brunsting, S., Pesch, U. and Feenstra, C.F.J. (2015) "How Stakeholder Interactions Can Reduce Space for Moral Considerations in Decision Making: A Contested CCS Project in the Netherlands," *Environment and Planning A* 47: 1963–1978.

Daukas, N. (2006) "Epistemic Trust and Social Location," *Episteme* 3(1–2): 109–124.

DesAutels, P. (2004) "Moral Mindfulness," in P. DesAutels and M. Urban Walker (eds.), *Moral Psychology: Feminist Ethics and Social Theory*, Lanham, MD: Rowman & Littlefield.

Dezecache, G. and Dunbar, R. (2012) "Sharing the Joke: The Size of Natural Language Groups," *Evolution & Human Behavior* 33(6): 775–779.

Dunbar, R. (1992) "Neocortex Size as a Constraint on Group Size in Primates," *Journal of Human Evolution* 22(6): 469–493.

Dunbar, R. (1993) "Coevolution of Neocortical Size, Group Size and Language in Humans," *Behavioral and Brain Sciences* 16(4): 681–735.

Dunbar, R. (2005) "Gossip in Evolutionary Perspective," *Review of General Psychology* 8: 100–110.

Dunbar, R. (2012) "On the Evolutionary Function of Song and Dance," in N. Bannan (ed.), *Music, Language and Human Evolution*, Oxford: Oxford University Press.

Elder, G. and Clipp, E. (1988) "Wartime Losses and Social Bonding: Influences across 40 Years in Men's Lives," *Psychiatry* 51(2): 177–198.

Feenstra, C.F.J., Mikunda, T. and Brunsting, S. (2010) "What Happened in Barendrecht?! Case Study on the Planned Onshore Carbon Dioxide Storage in Barendrecht, the Netherlands." www.osti.gov/etdeweb/biblio/21360732

Feinberg, J. (1968) "Collective Responsibility," *Journal of Philosophy* 65: 674–688.

Goldberg, S.(2010) *Relying on Others: An Essay in Epistemology*, Oxford:Oxford University Press.

Govier, T. (1993) "Self-Trust, Autonomy, and Self-Esteem," *Hypatia* 8(1): 99–120.

Grasswick, H. (2017) "Feminist Responsibilism, Situationism, and the Complexities of the Virtue of Trustworthiness," in A. Fairweather and M. Alfano (eds.), *Epistemic Situationism*, Oxford: Oxford University Press.

Gross, C. (2007) "Community Perspectives of Wind Energy in Australia: The Application of a Justice and Community Fairness Framework to Increase Social Acceptance," *Energy Policy* 35: 2727–2736.

Haas, Z., Halpern, J. and Li, L. (2006) "Gossip-Based ad hoc Routing," *IEEE/ACM Transactions on Networking (TON)* 14(3): 479–491.

Habermas, J. (1990) *Moral Consciousness and Communicative Action*, C. Lenhardt and S.W. Nicholsen (trans.), Cambridge, MA: MIT Press.

Hardin, R. (2002) *Trust and Trustworthiness*, New York: Russell Sage Foundation.

Harding, S. (1986) *The Science Question in Feminism*, Ithaca, NY: Cornell University Press.

Hobbes, T. ([1668] 1994) *Leviathan*, E. Curley (ed.), Indianapolis, IN: Hackett.

Huijts, N.M.A., Midden, C. and Meijnders, A.L. (2007) "Social Acceptance of Carbon Dioxide Storage," *Energy Policy* 35: 2780–2789.

Huijts, N.M.A., Molin, E.J.E. and Steg, L. (2012) "Psychological Factors Influencing Sustainable Energy Technology Acceptance: A Review-based Comprehensive Framework," *Renewable and Sustainable Energy Reviews* 16(1), 525–531.

Huijts, N.M.A., Molin, E.J.E. and Van Wee, B. (2014) "Hydrogen Fuel Station Acceptance: A Structural Equation Model Based on the Technology Acceptance Framework," *Journal of Environmental Psychology* 38, 153–166.

Jones, K. (1999) "Second-Hand Moral Knowledge," *Journal of Philosophy* 96(2): 55–78.

Jones, K. (2012) "Trustworthiness," *Ethics* 123(1): 61–85.

Krishnamurthy, M. (2015) "(White) Tyranny and the Democratic Value of Distrust," *The Monist* 98 (4): 391–406.

Kyne, D. and Bolin, B. (2016) "Emerging Environmental Justice Issues in Nuclear Power and Radioactive Contamination," *International Journal of Environmental Research and Public Health* 13(7): 700.

L'Orange Seigo, S., Dohle, S. and Siegrist, M. (2014) "Public Perception of Carbon Capture and Storage (CCS): A Review," *Renewable and Sustainable Energy Reviews* 38: 848–863.

McGeer, V. (2008) "Trust, Hope, and Empowerment," *Australasian Journal of Philosophy* 86(2): 237–254.

Meijnders, A., Midden, C., Olofsson, A., Öhman, S., Matthes, J., Bondarenko, O., Gutteling, J. and Rusanen, M. (2009) "The Role of Similarity Cues in the Development of Trust in Sources of Information about GM Food," *Risk Analysis*, 29(8): 1116–1128.

Milgram, S. (1967) "The Small World Problem," *Psychology Today* 1(1): 60–67.

Nietzsche, F. (1874 / 1997) *Untimely Meditations*, R.J. Hollingdale (trans.), D. Breazeale (ed.), Cambridge: Cambridge University Press.

Pettit, P. (1995) "The Cunning of Trust," *Philosophy and Public Affairs* 24(3): 202–225.

Plumer, B. (2016, November 29) "The Battle over the Dakota Access Pipeline, Explained." Vox. www.vox.com/2016/9/9/12862958/dakota-access-pipeline-fight

Porsius, J.T., Claassen, L., Weijland, P.E. and Timmermands, D.R.T. (2016) "'They Give You Lots of Information, but Ignore What It's Really About': Residents Experiences with the Planned Introduction of a New High-Voltage Power Line," *Journal of Environmental Planning and Management* 59(8): 1495–1512.

Potter, N. (2002) *How Can I be Trusted? A Virtue Theory of Trustworthiness*, Lanham, MD: Rowman & Littlefield.

Siegrist, M., Cvetkovich, G. and Roth, C. (2000) "Salient Value Similarity, Social Trust, and Risk/Benefit Perception," *Risk Analysis* 20(3), 353–362.

Surowiecki, J. (2005) *The Wisdom of Crowds*, New York: Anchor.

Ter Mors, E., Weenig, M.W.H., Ellemers, N. and Daamen, D.D.L. (2010) "Effective Communication about Complex Environmental Issues: Perceived Quality of Information about Carbon Dioxide Capture and Storage (CCS) Depends on Stakeholder Collaboration," *Journal of Environmental Psychology* 30: 347–357.

Terwel, B.W., Harinck, F., Ellemers, N. and Daamen, D.D.L. (2009) "How Organizational Motives and Communications Affect Public Trust in Organizations: The Case of Carbon Dioxide Capture and Storage," *Journal of Environmental Psychology* 29(2): 290–299.

Terwel, B.W., Harinck, F., Ellemers, N., and Daamen, D.D.L. (2010) "Voice in Political Decision-Making: The Effect of Group Voice on Perceived Trustworthiness of Decision Makers and Subsequent Acceptance of Decisions," *Journal of Experimental Psychology: Applied* 16(2): 173–186.

Townley, C. and Garfield, J. (2013) "Public Trust," in P. Makela and C. Townley (eds.), *Trust: Analytic and Applied Perspectives*, Amsterdam: Rodopi Press.

Wigley, D. and Shrader-Frechette, K. (1996) "Environmental Justice: A Louisiana Case Study," *Journal of Agricultural and Environmental Ethics* 9(1): 61–82.

Wylie, A. (2003) "Why Standpoint Matters," in R. Figueroa and S. Harding (eds.), *Science and Other Cultures: Issues in Philosophies of Science and Technology*, London: Routledge.

Further Reading

Alfano, M. (2016a) "The Topology of Communities of Trust," *Russian Sociological Review* 15(4): 30–56.

Baier, A. (1986) "Trust and Antitrust," *Ethics* 96: 231–160.

Bergmans, A. (2008) "Meaningful Communication among Experts and Affected Citizens on Risk: Challenge or Impossibility?" *Journal of Risk Research* 11(1–2): 175–193.

Dunbar, R. (2005) "Gossip in Evolutionary Perspective," *Review of General Psychology* 8: 100–110.

Grasswick, H. (2017) "Feminist Responsibilism, Situationism, and the Complexities of the Virtue of Trustworthiness," in A. Fairweather and M. Alfano (eds.), *Epistemic Situationism*, Oxford: Oxford University Press.

Habermas, J. (1990) *Moral Consciousness and Communicative Action*, C. Lenhardt and S.W. Nicholsen (trans.), Cambridge, MA: MIT Press.

Jones, K. (2012) "Trustworthiness," *Ethics* 123(1): 61–85.

Krishnamurthy, M. (2015) "(White) Tyranny and the Democratic Value of Distrust," *The Monist*, 98(4): 391–406.