Stanford Encyclopedia of Philosophy

Trust

First published Mon Feb 20, 2006; substantive revision Mon Aug 10, 2020

Trust is important, but it is also dangerous. It is important because it allows us to depend on others—for love, for advice, for help with our plumbing, or what have you—especially when we know that no outside force compels them to give us these things. But trust also involves the risk that people we trust will not pull through for us, since if there were some guarantee they would pull through, then we would have no need to trust them.[1] Trust is therefore dangerous. What we risk while trusting is the loss of valuable things that we entrust to others, including our self-respect perhaps, which can be shattered by the betrayal of our trust.

Because trust is risky, the question of when it is warranted is of particular importance. In this context, "warranted" means justified or well-grounded meaning, respectively, that the trust is rational (e.g. it is based on good evidence) or that it successfully targets a trustworthy person. If trust is warranted in these senses, then the danger of it is either minimized as with justified trust or eliminated altogether as with well-grounded trust. Leaving the danger of trust aside, one could also ask whether trust is warranted in the sense of being plausible. Trust may not be warranted in a particular situation because it is simply not plausible; the conditions necessary for it do not exist, as is the case when people feel only antagonism toward one another. This entry on trust is framed as a response to the general question of when trust is warranted, where "warranted" is broadly construed to include "justified," "well-grounded" and "plausible."

A complete philosophical answer to the above question must explore the various philosophical dimensions of trust, including the conceptual nature of trust and trustworthiness, the epistemology of trust, the value of trust, and the kind of mental attitude trust is. To illustrate how each of these matters is relevant, note that trust is warranted, that is,

  i. plausible, again, only if the conditions required for trust exist (e.g. some optimism about one another's ability). Knowing what these conditions are requires understanding the nature of trust.
  ii. well-grounded, only if the trustee (the one trusted) is trustworthy, which makes the nature of trustworthiness important in determining when trust is warranted.
 iii. justified, sometimes when the trustee is not in fact trustworthy, which suggests that the epistemology of trust is relevant.
 iv. justified, often because some value will emerge from the trust or because it is valuable in and of itself. Hence, the value of trust is important.
  v. plausible, only when it is possible for one to develop trust, given one's circumstances and the sort of mental attitude trust is. For instance, trust may not be the sort of attitude that one can will oneself to have without any evidence of a person's trustworthiness.

This piece explores these different philosophical issues about trust. It deals predominantly with interpersonal trust, which I take to be the dominant paradigm of trust. Although some philosophers write about trust that is not interpersonal, including trust in groups (Hawley 2017), institutional trust (i.e. trust in institutions; see e.g. Potter 2002, Govier 1997, Townley and Garfield 2013), trust in government (e.g. Hardin 2002, Budnik 2018) or science (e.g. Oreskes 2019), self-trust (Govier 1993, Lehrer 1997, Foley 2001, McLeod 2002, Goering 2009, Jones 2012b, Potter 2013), and trust in robots (e.g. Coeckelbergh

2012, Sullins 2020), most would agree that these forms of "trust" are coherent only if they share important features of (i.e. can be modeled on) interpersonal trust. I therefore assume that the dominant paradigm is interpersonal.

In addition, while my focus is mainly on trust and trustworthiness, I also discuss distrust (more so in this version of the entry than in previous versions). Distrust has received surprisingly little attention from philosophers, although it has recently become a topic of serious concern for some of them, particularly those who are interested in the politics of trust and distrust in societies marked by oppression and privilege. Relevant issues include when distrust is warranted by people who experience oppression and how misplaced distrust (i.e. in the oppressed) can be overcome by people who are privileged. I delve into these matters and also summarize the few theories that exist about the nature of distrust.

1. The Nature of Trust and Trustworthiness

Trust is an attitude that we have towards people whom we hope will be trustworthy, where trustworthiness is a property not an attitude. Trust and trustworthiness are therefore distinct although, ideally, those whom we trust will be trustworthy, and those who are trustworthy will be trusted. For trust to be plausible in a relationship, the parties to the relationship must have attitudes toward one another that permit trust. Moreover, for trust to be well-grounded, both parties must be trustworthy. (Note that here and throughout —unless specified otherwise—"trustworthiness" is understood in a thin sense according to which X is trustworthy for me just in case I can trust X.)

Trusting requires that we can, 1) be vulnerable to others—vulnerable to betrayal in particular; 2) rely on others to be competent to do what we wish to trust them to do; and 3) rely on them to be willing to do it. [2] Notice that the second two conditions refer to a connection between trust and reliance. For most philosophers, trust is a kind of reliance but is not *mere* reliance (Goldberg 2020). Rather, trust involves reliance "plus some extra factor" (Hawley 2014, 5). Controversy surrounds this extra factor, which generally concerns why the trustor (i.e. the one trusting) would rely on the trustee to be willing to do what they are trusted to do.

Trustworthiness is likewise a kind of reliability, although it's not obvious what kind. Clear conditions for trustworthiness are that the trustworthy person is competent and willing to do what they are trusted to do. Yet this person may also have to be willing for certain reasons or as a result of having a certain kind of

motive for acting (e.g. they care about the trustor).

This section explains these various conditions for trust and trustworthiness and highlights the controversy that surrounds the condition about motive and relatedly how trust differs from mere reliance. Included at the end is some discussion about the nature of distrust.

Let me begin with the idea that the trustor must accept some level of vulnerability or risk (Becker 1996; Baier 1986). Minimally, what this person risks, or is vulnerable to, is the failure by the trustee to do what the trustor is depending on them to do. The trustor might try to reduce this risk by monitoring or imposing certain constraints on the behavior of the trustee; but after a certain threshold perhaps, the more monitoring and constraining they do, the less they trust this person. Trust is relevant "before one can monitor the actions of … others" (Dasgupta 1988, 51) or when out of respect for others one refuses to monitor them. One must be content with them having some discretionary power or freedom, and as a result, with being somewhat vulnerable to them (Baier 1986; Dasgupta 1988).

One might think that if one is relying while trusting—that is, if trust is a species of reliance—then accepted vulnerability would not be essential for trust. Do we not rely on things only when we believe they will actually happen? And if we believe that, then we don't perceive ourselves as being vulnerable. Many philosophers writing on trust and reliance say otherwise. They endorse the view of Richard Holton, who writes, "When I rely on something happening … I [only] need to plan on it happening; I need to work around the supposition that it will [happen]" (Holton 1994, 3). I need not be certain of it happening, in other words, and I could even have doubts that it will happen (Goldberg 2020). I could therefore accept that I am vulnerable. I could do that while trusting if trust is a form of reliance.

What does trusting make us vulnerable to, in particular? Annette Baier writes that "trusting can be betrayed, or at least let down, and not just disappointed" (1986, 235). In her view, disappointment is the appropriate response when one merely relied on someone to do something but did not trust them to do it. To elaborate, although people who monitor and constrain others' behavior may rely on them, they do not trust them if their reliance can only be disappointed rather than betrayed. One can rely on inanimate objects, such as alarm clocks, but when they break, one is not betrayed although one may be disappointed. This point reveals that reliance without the possibility of betrayal (or at least "let down") is not trust; people who rely on one another in a way that makes this reaction impossible do not trust one another.

But does trust always involve the potential for betrayal? "Therapeutic trust" may be an exception (Nickel 2007, 318). (And for further exceptions, see e.g. Hinchman 2017.) To illustrate this type of trust, consider parents who "trust their teenagers with the house or the family car, believing that their [children] may well abuse their trust, but hoping by such trust to elicit, in the fullness of time, more responsible and responsive trustworthy behaviour" (McGeer 2008, 241, her emphasis; see also Horsburgh 1960 and Pettit 1995). Therapeutic trust is not likely to be betrayed rather than merely be disappointed. It is unusual in this respect (arguably) and in other respects that will become evident later on in this entry. The rest of this section deals with usual rather than unusual forms of trust and trustworthiness.

Without relying on people to display some competence, we also can't trust them. We usually trust people to do certain things, such as look after our children, give us advice, or be honest with us, which we wouldn't do that if we thought they lacked the relevant skills, including potentially moral skills of knowing what it means to be honest or caring (Jones 1996, 7). Rarely do we trust people completely (i.e. A simply trusts B). Instead, "trust is generally a three-part relation: A trusts B to do X" (Hardin 2002, 9)— or "A trusts B with valued item C" (Baier 1986) or A trusts B in domain D (D'Cruz 2019, Jones 2019).[3] To have trust in a relationship, we do not need to assume that the other person will be competent in every way. Optimism about the person's competence in at least one area is essential, however.

When we trust people, we rely on them not only to be competent to do what we trust them to do, but also to be willing or motivated to do it. We could talk about this matter either in terms of what the trustor expects of the trustee or in terms of what the trustee possesses: that is, as a condition for trust or as a condition for trustworthiness (and the same is true, of course, of the competence condition). For simplicity's sake and to focus some of this section on trustworthiness rather than trust, I'll refer to the motivation of the trustee mostly as a condition for trustworthiness.

Although both the competence and motivational elements of trustworthiness are crucial, the exact nature of the latter is unclear. For some philosophers, it matters only that the trustee is motivated, where the central problem of trustworthiness in their view concerns the probability that this motivation will exist or endure (see, e.g., Hardin 2002, 28; Gambetta 1988). Jones calls these "risk-assessment views" about trust (1999, 68). According to them, we trust people whenever we perceive that the risk of relying on them to act a certain way is low and so we rely on (i.e., "trust") them. They are trustworthy if they are willing, for whatever reason, to do what they are trusted to do. Risk-assessment theories make no attempt to distinguish between trust and mere reliance and have been criticized for this reason (see e.g. Jones 1999).

By contrast, other philosophers say that just being motivated to act in the relevant way is not sufficient for trustworthiness; according to them, the nature of the motivation matters, not just its existence or duration. It matters in particular, they say, for explaining the trust-reliance distinction, which is something they aim to do. The central problem of trustworthiness for them is not simply whether but also how the trustee is motivated to act. Will that person have the kind of motivation that makes trust appropriate? Katherine Hawley identifies theories that respond to this question as "motives-based" theories (2014).

To complicate matters, there are "non-motives-based theories," which are also not risk-assessment theories (Hawley 2014). They strive to distinguish between trust and mere reliance, though not by associating a particular kind of motive with trustworthiness. Since most philosophical debate about the nature of trust and trustworthiness centers on theories that are either motives-based or non-motives-based, let me expand on these categories.

1.1 Motives-based theories

Philosophers who endorse this type of theory differ in terms of what kind of motive they associate with trustworthiness. For some, it is self-interest, while for others, it is goodwill or an explicitly moral motive, such as moral integrity or virtue.[4]

For example, Russell Hardin defines trustworthiness in terms of self-interest in his "encapsulated interests" account (2002). He says that trustworthy people are motivated by their own interest to maintain the relationship they have with the trustor, which in turn encourages them to encapsulate the interests of that person in their own interests. In addition, trusting people is appropriate when we can reasonably expect them to encapsulate our interests in their own, an expectation which is missing with mere reliance.

Hardin's theory may be valuable in explaining many different types of trust relationships, including those between people who can predict little about one another's motives beyond where their self-interest lies. Still, it is problematic. To see why, consider how it applies to a sexist employer who has an interest in maintaining relationships with women employees, who treats them reasonably well as a result, but whose interest stems from a desire to keep them around so that he can daydream about having sex with them. This interest conflicts with an interest the women have in not being objectified by their employer. At the same time, if they were not aware of his daydreaming—say they are not—then he can ignore this particular interest of theirs. He can encapsulate enough of their other interests in his own to keep his relationships with them going. And this would make him trustworthy on Hardin's account. But is he

trustworthy? I assume the answer is "no" or that at least the women themselves would say "no" if they knew the main reason for their employment. The point is that being motivated by a desire to maintain a relationship (the central motivation of a trustworthy person on the encapsulated interests view) may not require one to adopt all of the interests of the trustor that would actually make one trustworthy to that person. In the end, the encapsulated interests view might describe only reliability, not trustworthiness. The sexist employer may reliably treat the women well, because of his interest in daydreaming about them, but he is not trustworthy because of why he treats them well.

A different type of theory is what Jones calls a "will-based" account, which finds trustworthiness only where the trustee is motivated by goodwill (Jones 1999, 68). This view originates in the work of Annette Baier and is influential, even outside of moral philosophy (e.g. in bioethics and law, especially fiduciary law; see e.g. Pellegrino and Thomasma 1993, O'Neill 2002, and Fox-Decent 2005). According to it, a trustee who is trustworthy will act out of goodwill toward the trustor, to what or to whom the trustee is entrusted with, or both. While many readers might find the goodwill view problematic—surely we can trust people without presuming their goodwill!—it is immune to a criticism that applies to Hardin's theory and also to risk-assessment theories. The criticism is that they fail to require that the trustworthy person care about (i.e. feel goodwill towards) the trustor, or care about what the trustor cares about. As we have seen, such caring appears to be central to a complete account of trustworthiness.

The particular reason why care may be central is that it allows us to grasp how trust and reliance differ. Earlier I suggested that they differ because only trust can be betrayed (or at least let down). But why is that true? Why can trust be betrayed, while mere reliance can only be disappointed? The answer Baier gives is that betrayal is the appropriate response to someone on whom one relied to act out of goodwill, as opposed to ill will, selfishness, or habit bred out of indifference (1986, 234–5; see also Baier 1991). Those who say that trusting could involve relying on people to act on a motive like ill will or selfishness will have trouble distinguishing between trust and mere reliance.

While useful in some respects, Baier's will-based account is not perfect. Criticisms have been made that suggest goodwill is neither necessary nor sufficient for trustworthiness. It is not necessary because we can trust other people without presuming that they have goodwill (e.g. O'Neill 2002; Jones 2004), as we arguably do when we put our trust in strangers.

As well as being unnecessary, goodwill may not be sufficient for trustworthiness, and that is true for three reasons. First, someone trying to manipulate you—a "confidence trickster" (Baier 1986)—could "rely on your goodwill without trusting you," say, to give them money (Holton 1994, 65). You are not trustworthy for them, despite your goodwill, because they are not trusting you but rather are just trying to trick you. Second, basing trustworthiness on goodwill alone cannot explain unwelcome trust. We do not always welcome people's trust, because trust can be burdensome or inappropriate. When that happens, we object not to their optimism about our goodwill, however (who would object to that?), but only to the fact that they are counting on us. Third, we can expect people to be reliably benevolent toward us without trusting them (Jones 1996, 10). We can think that their benevolence is not shaped by the sorts of values that for us are essential to trustworthiness.[5]

Criticisms about goodwill not being sufficient for trustworthiness have prompted revisions to Baier's theory and occasionally to the development of new will-based theories. For example, in response to the first criticism—about the confidence trickster—Zac Cogley argues that trust involves the belief not simply that the trustee will display goodwill toward us but that this person owes us goodwill (2012). Since the confidence trickster doesn't believe that their mark owes them goodwill, they don't trust this person, and neither is this person trustworthy for them. In response to the second criticism—the one about unwelcome trust—Jones claims that optimism about the trustee's goodwill must be coupled with the expectation that the trustee will be "favorably moved by the thought that [we are] counting on her" (1996, 9). Jones does

that in her early work on trust where she endorses a will-based theory. Finally, in response to the third concern about goodwill not being informed by the sorts of values that would make people trustworthy for us, some maintain that trust involves an expectation about some shared values, norms, or interests (Lahno 2001, 2020; McLeod 2002, 2020; Mullin 2005; Smith 2008). (To be clear, this last expectation tends not to be combined with goodwill to yield a new will-based theory or theories.)

One final criticism of will-based accounts concerns how "goodwill" should be interpreted. In much of the discussion above, it is narrowly conceived so that it involves friendly feeling or personal liking. Yet Jones urges us in her early work on trust to understand goodwill so that it could be grounded in benevolence, conscientiousness, or the like, *or* in friendly feeling (1996, 7). Then in her later work, however, she worries that by defining goodwill so broadly we "turn it into a meaningless catchall that merely reports the presence of some positive motive, and one that may or may not even be directed toward the truster" (2012a, 67). Jones abandons her own will-based theory upon rejecting both a narrow and a broad construal of goodwill. If her concerns about defining goodwill are valid, then will-based theories are in serious trouble.

To recapitulate about encapsulated-interest and will-based theories, they claim that a trustworthy person is motivated by self-interest or goodwill, respectively. Encapsulated-interest theories struggle to explain how trustworthiness differs from mere reliability, while will-based theories are faced with the criticism that goodwill is neither necessary nor sufficient for trustworthiness. Some philosophers who say that goodwill is insufficient develop alternative will-based theories. An example is Cogley's theory according to which trust involves a normative expectation of goodwill (2012).

The field of motives-based theories is not exhausted by encapsulated-interest and will-based theories, however. Other motives-based theories include those that describe the motive of trustworthy people in terms of a moral commitment, moral obligation, or virtue. To expand, consider that one could make sense of the trustworthiness of a stranger by presuming that the stranger is motivated not by self-interest or goodwill, but by a commitment to stand by their moral values. In that case, I could trust a stranger to be decent simply by presuming that she is committed to common decency. Ultimately, what I am presuming about the stranger is moral integrity, which some say is the relevant motive for trust relations (those that are prototypical; see McLeod 2002). Others identify this motive similarly as moral obligation, and say it is ascribed to the trustee by the very act of trusting them (Nickel 2007; for a similar account, see Cohen and Dienhart 2013). Though compelling in some respects, the worry about these theories is that they moralize trust inappropriately by demanding that the trustworthy person have a moral motive (see e.g. Mullin 2005, Jones 2017).

Yet one might insist that it is appropriate to moralize trust or at least moralize trustworthiness, which we tend to think of as a virtuous character trait. Nancy Nyquist Potter refers to the trait as "full trustworthiness," and distinguishes it from "specific trustworthiness": trustworthiness that is specific to certain relationships (and equivalent to the thin sense of trustworthiness I use throughout; 2002, 25). To be fully trustworthy, one must have a disposition to be trustworthy toward everyone, according to Potter. Call this the "virtue" account.

It may sound odd to insist that trustworthiness is a virtue or, in other words, a moral disposition to be trustworthy (Potter 2002, 25; Hardin 2002, 32). What disposition exactly is it meant to be? A disposition normally to honor people's trust? That would be strange, since trust can be unwanted if the trust is immoral (e.g., being trusted to hide a murder) or if it misinterprets the nature of one's relationship with the trustee (e.g., being trusted to be friends with a mere acquaintance). Perhaps trustworthiness is instead a disposition to respond to trust in appropriate ways, given "who one is in relation" to the trustor and given other virtues that one possesses or ought to possess (e.g., justice, compassion) (Potter 2002, 25). This is essentially Potter's view. Modeling trustworthiness on an Aristotelian conception of virtue, she defines a

trustworthy person as "one who can be counted on, as a matter of the sort of person he or she is, to take care of those things that others entrust to one and (following the Doctrine of the Mean) whose ways of caring are neither excessive nor deficient" (her emphasis; 16).[6] A similar account of trustworthiness as a virtue—an epistemic one, specifically—can be found in the literature on testimony (see Frost-Arnold 2014; Daukas 2006, 2011).

Criticism of the virtue account comes from Karen Jones (2012a). As she explains, if being trustworthy were a virtue, then being untrustworthy would be a vice, but that can't be right because we can never be required to exhibit a vice, yet we can be required to be untrustworthy (84). An example occurs when we are counted on by two different people to do two incompatible things and being trustworthy to the one demands that we are untrustworthy to the other (83). To defend her virtue theory, Potter would have to insist that in such situations, we are required either to disappoint someone's trust rather than be untrustworthy, or to be untrustworthy in a specific not a full sense.[7]

Rather than cling to a virtue theory, however, why not just endorse the thin conception of trustworthiness (i.e. "specific trustworthiness"), according to which X is trustworthy for me just in case I can trust X? Two things can be said. First, the thick conception—of trustworthiness as a virtue—is not meant to displace the thin one. We can and do refer to some people as trustworthy in the specific or thin sense and others as trustworthy in the full or thick sense. Second, one could argue that the thick conception explains better than the thin one why fully trustworthy people are as dependable as they are. It is ingrained in their character. They therefore must have an ongoing commitment, and better still, their commitment must come from a source that is compatible with trustworthiness (i.e. virtue as opposed to mere self-interest).

An account of trustworthiness that includes the idea that trustworthiness is a virtue will seem ideal only if we think that the genesis of the trustworthy person's commitment matters. If we believe—like risk-assessment theorists—that it matters only whether, not how, the trustor will be motivated to act, then we could assume that ill will can do the job as well as a moral disposition. Such controversy explains how and why motives-based and risk-assessment theories diverge from one another.

1.2 Non-motives-based theories

A final category are theories that base trustworthiness neither on the kind of motivation a trustworthy person has nor on the mere willingness of this person to do what they are relied on to do. These are non-motives-based and also non-risk-assessment theories. The conditions that give rise to trustworthiness according to them reside ultimately in the stance the trustor takes toward the trustee or in what the trustor believes they ought to be able to expect from this person (so in normative expectations of them). These theories share with motives-based theories the goal of describing how trust differs from mere reliance.

An example is Richard Holton's theory of trust (1994). Holton argues that trust is unique because of the stance the trustor takes toward the trustee: the "participant stance," which involves treating the trustee as a person—someone who is responsible for their actions—rather than simply as an object (see also Strawson 1974). In the case of trust specifically, the stance entails a readiness to feel betrayal (Holton 1994, 4). Holton's claim is that this stance and this readiness are absent when we merely rely on someone or something.

Although Holton's theory has garnered positive attention (e.g. by Hieronymi 2008, McGeer 2008), some find it dissatisfying. For example, some argue that it does not obviously explain what would justify a reaction of betrayal, rather than mere disappointment, when someone fails to do what they are trusted to do (Jones 2004; Nickel 2007). They could fail to do it just by accident, in which case feelings of betrayal would be inappropriate (Jones 2004). Others assert, by contrast, that taking the participant stance toward

someone "does not always mean trusting that person: some interactions [of this sort] lie outside the realm of trust and distrust" (Hawley 2014, 7). To use an example from Hawley, my partner could come to rely on me to make him dinner every night in a way that involves him taking the participant stance toward me. But he needn't trust me to make him dinner and so needn't feel betrayed if I do not. He might know that I am loath for him to trust me in this regard: that is, "to make this a matter of trust" between us (Hawley 2014, 7).

Some philosophers have expanded on Holton's theory in a way that may deflect some criticism of it. Margaret Urban Walker emphasizes that in taking a participant stance, we hold people responsible (2006, 79). We expect them to act not simply as we assume they *will*, but as they *should*. We have, in other words, normative rather than merely predictive expectations of them. Call this a "normative-expectation" theory, which again is an elaboration on the participant-stance theory. Endorsed by Walker and others (e.g. Jones 2004 and 2012a; Frost-Arnold 2014), this view explains the trust-reliance distinction in terms of the distinction between normative and predictive expectations. It also describes the potential for betrayal in terms of the failure to live up a normative expectation.

Walker's theory is non-motives-based because it doesn't specify that trustworthy people must have a certain kind of motive for acting. She says that trustworthiness is compatible with having many different kinds of motives, including, among others, goodwill, "pride in one's role," "fear of penalties for poor performance," and "an impersonal sense of obligation" (2006, 77). What accounts for whether someone is trustworthy in her view is whether they act as they should, not whether they are motivated in a certain way. (By contrast, Cogley's normative-expectation theory says that the trustworthy person both will and ought to act with goodwill. His theory is motives-based.)

Prominent in the literature is a kind of normative-expectation theory called a "trust- (or dependence-) responsive" theory (see e.g. Faulkner and Simpson 2017, 8; Faulkner 2011, 2017; Jones 2012a, 2017, 2019; McGeer and Petit 2017). According to this view, being trustworthy involves being appropriately responsive to the reason you have to do X—what you are being relied on (or "counted on"; Jones 2012a) to do—when it's clear that someone is in fact relying on you. The reason you have to do X exists simply because someone is counting on you; you should do it, other things being equal, for this reason. Being appropriately responsive to it, moreover, just means that you find it compelling (Jones 2012a, 70–71). The person trusting you expects you to have this reaction; in other words, they have a normative expectation that the "manifest fact of [their] reliance will weigh on you as a reason for choosing voluntarily to X" (McGeer and Petit 2017, 16). This expectation is missing in cases of mere reliance. When I merely rely on you, I do not expect this fact (i.e. my reliance) to weigh on you as I do when I trust you.

Although trust-responsiveness theories might seem motives-based, they are not. One might think that they require, in order to be trustworthy, that you be motivated by the fact that you are being counted on. Instead, they demand only that you be appropriately responsive to the reason you have to do what you are being depended on to do. As Jones explains, you could be responsive in this way and act ultimately out of goodwill, conscientiousness, love, duty, or the like (2012a, 66). The reaction I expect of you, as the trustor, is compatible with you acting on different kinds of motives, although to be clear, some motives are ruled out, including indifference and ill will (Jones 2012a, 68). Being indifferent or hateful towards me means that you are unlikely to view me counting on you as a reason to act. Hence, if I knew you were indifferent or hateful, I would not expect you to be trust responsive.

Trust-responsiveness theories are less restrictive than motives-based theories when it comes to defining what motives you need to be trustworthy. At the same time, they are more restrictive when it comes to stating whether, in order to be trustworthy or trusted, you must be aware that you are being counted on. You couldn't be trust responsive otherwise; and so, in trusting you, I must "make clear to you my assumption that you will prove reliable in doing X" (McGeer and Pettit 2017, 16). I do not have to do that

if, in trusting you, I am relying on you instead to act with a motive like goodwill. Baier herself allows that trust can exist where the trustee is unaware of it (1986, 235; see also Hawley 2014; Lahno 2020). For her, trust is ubiquitous (Jones 2017, 102), in part for this reason; we trust people in a myriad of different ways every single day, often without them knowing it. If she's right about this fact, then trust-responsiveness theories are incomplete.

These theories are also vulnerable to objections raised against normative-expectation theories, because they are again a type of normative-expectation theory. One such concern comes from Hawley. In writing about both trust and distrust, she states that "we need a story about when trust, distrust or neither is objectively appropriate—what is the worldly situation to which (dis)trust] is an appropriate response? When is it appropriate to have (dis)trust-related normative expectations of someone?" (2014, 11). Normative-expectation theories tend not to provide an answer. And trust-responsiveness theories suggest only that trust-related normative expectations are appropriate when certain motives are absent (e.g., ill will), which may not to be enough.

Hawley responds to the above concern with her "commitment account" of trust (2014, 2019). This theory states that in trusting others, we believe they have a commitment to doing what we are trusting them to do (2014, 10), a fact which explains why we expect them to act this way, and also why we fail to do so in cases like that of my partner relying on me to make dinner; he knows I have no commitment to making his dinner (or anyone else's) repeatedly. For Hawley, the relevant commitments "can be implicit or explicit, weighty or trivial, conferred by roles and external circumstances, default or acquired, welcome or unwelcome" (11). They also needn't actually motivate the trustworthy person. Her theory is non-motives-based because it states that to "be trustworthy, in some specific respect, it is enough to behave in accordance with one's commitment, regardless of motive" (2014, 16). Similarly, to trust me to do something, it is enough to believe that I "have a commitment to do it, and that I will do it, without believing that I will do it *because* of my commitment" (2014, 16; her emphasis).

Notice that unlike trust-responsive theories, the commitment account does not require that the trustee be aware of the trust in order to be trustworthy. This person simply needs to have a commitment and to act accordingly. They don't even need to be committed to the trustor, but rather could be committed to anyone and one could trust them to follow through (Hawley 2014, 11). Thus, relying on a promise your daughter's friend makes *to her* to take her home from the party would count as an instance of trust (Hawley 2014, 11). In this way, the commitment account is less restrictive than trust-responsive theories are. In being non-motives-based, Hawley's theory is also less restrictive than any motives-based theory. Trust could truly be ubiquitous if she's correct about the nature of it.

Like the other theories considered here, however, the commitment account is open to criticisms. One might ask whether Hawley gives a satisfactory answer to the question that motivates her theory: when can we reasonably have the normative expectations of someone that go along with trusting them? Hawley's answer is, when this person has the appropriate commitment, where "commitment" is understood very broadly. Yet where the relevant commitment is implicit or unwelcome, it's not clear that we can predict much about the trustee's behavior. In cases like these, the commitment theory may have little to say about whether it is reasonable to trust.

A further criticism comes from Andrew Kirton (forthcoming) who claims that we sometimes trust people to act contrary to what they are committed to doing. His central example involves a navy veteran, an enlisted man, whose ship sunk at sea and who trusted those who rescued them (navy men) to ignore a commitment they had to save the officers first, because the officers were relatively safe on lifeboats compared to the enlisted men who were struggling in the water. Instead the rescuers adhered to their military duty, and the enlisted man felt betrayed by them for nearly letting him drown. If compelling, then this example shows that trust and commitment can come apart and that Hawley's theory is incomplete.[8]

The struggle to find a complete or unified theory of trust has led some philosophers to be pluralists about trust—that is, to say, "we must recognise plural forms of trust" (Simpson 2012, 551) or accept that trust is not just one form of reliance, but many forms of it (see also Jacoby 2011, Scheman 2020, McLeod 2020). Readers may be led to this conclusion from the rundown I've given of the many different theories of trust in philosophy and the objections that have been raised to them. Rather than go in the direction of pluralism, however, most philosophers continue to debate what unifies trust in a way that makes it different from mere reliance. They tend to believe that a unified and suitably developed motives-based theory or non-motives-based theory can explain this difference, although there is little consensus about this theory is like.

In spite of there being little settled agreement in philosophy about trust, there are thankfully things we can say for certain about it that are relevant to deciding when it is warranted. The trustor must be able to accept that by trusting, they are vulnerable, usually to betrayal. Also, the trustee must be competent and willing to do what the trustor expects of them and may have to be willing because of certain attitudes they have. Last, in cases of paradigmatic trust at least, the trustor must be able to rely on the trustee to exhibit this competence and willingness.

1.3 Distrust

As suggested above, distrust has been somewhat of an afterthought for philosophers (Hawley 2014),[9] although their attention to it has grown recently. As with trust and trustworthiness, philosophers would agree that distrust has certain features, although the few who have developed theories of distrust disagree ultimately about the nature of it.

The following are features of distrust that are relatively uncontroversial (see D'Cruz 2020):

1) Distrust is not just the absence of trust since it is possible to neither distrust nor trust someone (Hawley 2014, 3; Jones 1996, 16; Krishnamurthy 2015). There is gap between the two—"the possibility of being suspended between" them (Ullman-Margalit 2017, 184). (For disagreement, see Faulkner 2017.)

2) Although trust and distrust are not exhaustive, they are exclusive; one cannot at the same time trust and distrust someone about the same matter (Ullman-Margalit 2017, 201).

3) Distrust is "not mere nonreliance" (Hawley 2014, 3). I could choose not to rely on a colleague's assistance because I know she is terribly busy, not because I distrust her.

4) Relatedly, distrust has a normative dimension. If I distrusted a colleague for no good reason and they found out about it, then they would probably be hurt or angry. But the same reaction would not accompany them knowing that I decided not to rely on them (Hawley 2014). Being distrusted is a bad thing (Domenicucci and Holton 2017, 150; D'Cruz 2019, 935), while not being relied on needn't be bad at all.

5) Distrust is normally a kind of nonreliance, just as trust is a kind (or many kinds) of reliance. Distrust involves "action-tendencies" of avoidance or withdrawal (D'Cruz 2019, 935–937), which make it incompatible with reliance—or at least complete reliance. We can be forced to rely on people we distrust, yet even then, we try to keep them at as safe a distance as possible.

Given the relationship between trust and distrust and the similarities between them (e.g., one is "richer than [mere] reliance" and the other is "richer than mere nonreliance"; Hawley 2014, 3), one would think that any theory of trust should be able to explain distrust and vice versa. Hawley makes this point and criticizes theories of trust for not being able to make sense of distrust (2014, 6–9). For example, will-

based accounts imply that distrust must be nonreliance plus an expectation of ill will, yet the latter is not required for distrust. I could distrust someone because he is careless, not because he harbors ill will toward me (Hawley 2014, 6).

Hawley defends her commitment account of trust, in part, because she believes it is immune to the above criticism. It says that distrust is nonreliance plus the belief that the person distrusted is committed to doing what we will not rely on them to do. In spite of them being committed in this way (or so we believe), we do not rely on them (2014, 10). This account does not require that we impute any particular motive or feeling to the one distrusted, like ill will. At the same time, it tells us why distrust is not mere nonreliance and also why it is normative; the suspicion of the one distrusted is that they will fail to meet a commitment they have, which is bad.

Some have argued that Hawley's theory of distrust is subject to counterexamples (D'Cruz 2020; Tallant 2017). For example, Jason D'Cruz describes a financier who "buys insurance on credit defaults, positioning himself to profit when borrowers default" (2020, 45). The financier believes that the borrowers have a commitment not to default, and he does not rely on them to meet this commitment. The conclusion that Hawley's theory would have us reach is that he distrusts the borrowers, which doesn't seem right.

A different kind of theory of distrust can be found in the work of Meena Krishnamurthy (2015), who is interested specifically in the value that distrust has for political democracies, and for political minorities in particular (2015). She offers what she calls a "narrow normative" account of distrust that she derives from the political writings of Martin Luther King Jr. The account is narrow because it serves a specific purpose: of explaining how distrust can motivate people to resist tyranny. It is normative because it concerns what they ought to do (again, resist; 392). The theory says that distrust is the confident belief that others will not act justly. It needn't involve an expectation of ill will; King's own distrust of White moderates was not grounded in such an expectation (Krishnamurthy 2015, 394). To be distrusting, one simply has to believe that others will not act justly, whether out of fear, ignorance, or what have you.

D'Cruz complains that Krishnamurthy's theory is too narrow because it requires a belief that the one distrusted will fail to *do* something (i.e. act justly) (2020); but one can be distrustful of someone—say a salesperson who comes to your door (Jones 1996)—without predicting that they will do anything wrong or threatening. D'Cruz does not explain, however, why Krishnamurthy needs to account for cases like these in her theory, which again is meant to serve a specific purpose. Is it important that distrust can take a form other than "X distrusts Y to [do] Φ" for it to motivate political resistance (D'Cruz 2020, 45)? D'Cruz's objection is sound only if the answer is "yes."

Nevertheless, D'Cruz's work is helpful in showing what a descriptive account of distrust should look like —that is, an account that unlike Krishnamurthy's, tracks how we use the concept in many different circumstances. He himself endorses a normative-expectation theory, according to which distrust involves "a tendency to withdraw from reliance or vulnerability in contexts of normative expectation, based on a construal of a person or persons as malevolent, incompetent, or lacking integrity" (2019, 936). D'Cruz has yet to develop this theory fully, but once he does so, it will almost certainly be a welcome addition to the scant literature in philosophy on distrust.

In sum, among the relatively few philosophers who have written on distrust, there is settled agreement about some of its features but not about the nature of distrust in general. The agreed-upon features tell us something about when distrust is warranted (i.e. plausible). For distrust in someone to be plausible, one cannot also trust that person, and normally one will not be reliant on them either. Something else must be true as well, however. For example, one must believe that this person is committed to acting in a certain way but will not follow through on this commitment. The "something else" is crucial because distrust is

not the negation of trust and neither is it mere nonreliance.

Philosophers have said comparatively little about what distrust is, but a lot about how distrust tends to be influenced by negative social stereotypes that portray whole groups of people as untrustworthy (e.g. Potter 2020, Scheman 2020, D'Cruz 2019, Fricker 2007). Trusting attitudes are similar—who we trust can depend significantly on social stereotypes, positive ones—yet there is less discussion about this fact in the literature on trust. This issue concerns the rationality (more precisely, the *ir*rationality) of trust and distrust, which makes it relevant to the next section: on the epistemology of trust.

2. The Epistemology of Trust

Writings on this topic obviously bear on the issue of when trust is warranted (i.e. justified). The central epistemological question about trust is, "Ought I to trust or not?" That is, given the way things seem to me, is it reasonable for me to trust? People tend to ask this sort of question only in situations where they can't take trustworthiness for granted—that is, where they are conscious of the fact that trusting could get them into trouble. Examples are situations similar to those in which they have been betrayed in the past or unlike any they have ever been in before. The question, "Ought I to trust?" is therefore particularly pertinent to a somewhat odd mix of people that includes victims of infidelity, abuse, or the like, as well as immigrants and travelers.

The question "Ought I to distrust?" has received relatively little attention in philosophy despite it arguably being as important as the question of when to trust. People can get into serious trouble by distrusting when they ought to be trusting, not just by trusting when they ought to be distrusting. The harms of misplaced distrust are both moral and epistemic and include dishonoring people, being out of harmony with them, and being deprived of knowledge via testimony (D'Cruz 2019, Fricker 2007). Presumably because they believe that the harms of misplaced trust are greater (D'Cruz 2019), philosophers—and consequently I, in this entry—focus more on the rationality of trusting, as opposed to distrusting.

Philosophical work relevant to the issue of how to trust well appears either under the general heading of the epistemology or rationality of trust (e.g. Baker 1987, Webb 1992, Wanderer and Townsend 2013) or under the specific heading of testimony—that is, of putting one's trust in the testimony of others. This section focuses on the epistemology of trust generally rather than on trust in testimony specifically. There is a large literature on testimony (see the entry in this encyclopedia) and on the related topic of epistemic injustice, both of which I discuss only insofar as they overlap with the epistemology of trust.

Philosophers sometimes ask whether it could ever be rational to trust other people. This question arises for two reasons. First, it appears that trust and rational reflection (e.g. on whether one should be trusting) are in tension with one another. Since trust inherently involves risk, any attempt to eliminate that risk through rational reflection could eliminate one's trust by turning one's stance into mere reliance. Second, trust tends to give us blinkered vision: it makes us resistant to evidence that may contradict our optimism about the trustee (Baker 1987; Jones 1996 and 2019). For example, if I trust my brother not to harm anyone, I will resist the truth of any evidence to the contrary. Here, trust and rationality seem to come apart.

Even if some of our trust could be rational, one might insist that not all of it could be rational for various reasons. First, if Baier is right that trust is ubiquitous (1986, 234), then we could not possibly subject all of it to rational reflection. We certainly could not reflect on every bit of knowledge we've acquired through the testimony of others, such as that the earth is round or Antarctica exists (Webb 1993, Fricker 1995, Coady 1992). Second, bioethicists point out that some trust is unavoidable and occurs in the absence of rational reflection (e.g. trust in emergency room nurses and physicians; see Zaner 1991).

Lastly, some trust—namely the therapeutic variety—purposefully leaps beyond any evidence of trustworthiness in an effort to engender trustworthiness in the trustee. Is this sort of trust rational? Perhaps not, given that there isn't sufficient evidence for it.

Many philosophers respond to the skepticism about the rationality of trust by saying that rationality, when applied to trust, needs to be understood differently than it is in each of the skeptical points above. There, 'rationality' means something like this: it is rational to believe in something only if one has verified that it will happen or done as much as possible to verify it. For example, it is rational for me to believe that my brother has not harmed anyone only if the evidence points in that direction and I have discovered that to be the case. As we've seen, problems exist with applying this view of rationality to trust. Yet the rationality of trust needn't be understood according to it. This view is both "truth-directed" and "internalist," but the rationality of trust could be "end-directed" or "externalist." Alternatively, its rationality could be internalist without requiring that we have done the evidence gathering just discussed. Let me expand on these possibilities, starting with those that concern truth- or end-directed rationality.

2.1 Truth- vs. end-directed rationality

In discussing the rationality of trust, some authors distinguish between these two types of rationality (also referred to as epistemic vs. strategic rationality; see e.g. Baker 1987). One could say that we are rational in trusting emergency room physicians, for example, not necessarily because we have good reason to believe they are trustworthy (thus, our rationality is not truth-directed), but because by trusting them, we can remain calm in a situation over which we have little control (our rationality is therefore end-directed). Similarly, it may be rational for me to trust my brother not because I have good evidence of his trustworthiness but rather because trusting him is essential to our having a loving relationship.[10]

Trust can be rational, then, depending on whether one conceives of rationality as truth-directed or end-directed. Notice that it matters also how one conceives of trust. For example, if trust is a belief in someone's trustworthiness (see Trust and the Will), then whether the rationality of trust can be end-directed will depend on whether the rationality of a belief can be end-directed. To put the point more generally, how trust is rationally justified will depend on how beliefs are rationally justified (Jones 1996).

Some of the literature on trust and rationality concerns whether the rationality of trust can indeed be end-directed and also what could make therapeutic trust and the like rational. Pamela Hieronymi argues that the ends for which we trust cannot provide reasons for us to trust in the first place (2008). Considerations about how useful or valuable trust is do not bear on the truth of a trusting belief (i.e., a belief in someone's trustworthiness). But Hieronymi claims that trust, in a pure sense at least, always involves a trusting belief. How then does she account for trust that is motivated by how therapeutic (i.e. useful) the trust will be? She believes that trust of this sort is not pure or full-fledged trust. As she explains, people can legitimately complain about not being trusted fully when they are trusted in this way, which occurs when other people lack confidence in them but trust them nonetheless (230; see also Lahno 2001, 184–185).

By contrast, Victoria McGeer believes that trust is more substantial or pure when the available evidence does not support it (2008). She describes how trust of this sort—what she calls "substantial trust"—could be rational and does so without appealing to how important it might be or to the ends it might serve, but instead to whether the trustee will be trustworthy.[11] According to McGeer, what makes "substantial trust" rational is that it involves hope that the trustees will do what they are trusted to do, which "can have a galvanizing effect on how [they] see themselves, as trustors avowedly do, in the fullness of their potential" (252; see also McGeer and Pettit 2017). Rather than complain (as Hieronymi would assume that trustees might) about trustors being merely hopeful about their trustworthiness, they could respond well to the trustors' attitude toward them. Moreover, if it is likely that they will respond well—in other

words, they will be trust-responsive—then the trust in them must be epistemically rational. That is particularly true if being trustworthy involves being trust-responsive, as it does for McGeer (McGeer and Pettit 2017).

McGeer's work suggests that all trust—even therapeutic trust—can be rational in a truth-directed way. As we've seen, there is some dispute about whether trust can be rational in just an end-directed way. What matters here is whether trust is the sort of attitude whose rationality could be end-directed.

2.2 Internalism vs. externalism

Philosophers who agree that trust can be rational (in a truth- or end-directed way or both) tend to disagree about the extent to which reasons that make it rational must be accessible to the trustor. Some say that these reasons must be available to this person in order for their trust to be rational; in that case, the person is or could be internally justified in trusting as they do. Others say that the reasons need not be internal but can instead be external to the trustor and lie in what caused the trust, or, more specifically, in the epistemic reliability of what caused it. The trustor also needn't have access to or be aware of the reliability of these reasons. The latter's epistemology of trust is externalist, while the former's is internalist.

To elaborate, some epistemologists write as though trust is only rational if the trustor themselves has rationally estimated the likelihood that the trustee is trustworthy. For example, Russell Hardin implies that if my trust in you is rational, then "I make a rough estimate of the truth of [the] claim … that you will be trustworthy under certain conditions … and then I correct my estimate, or 'update,' as I obtain new evidence on you" (2002, 112). On this view, I must have reasons for my estimate or for my updates (Hardin 2002, 130), which could come from inductive generalizations I make about my past experience, from my knowledge that social constraints exist that will encourage your trustworthiness or what have you. Such an internalist epistemology of trust is valuable because it coheres with the commonsense idea that one ought to have good reasons for trusting people—reasons grounded in evidence that they will be trustworthy—particularly when something important is at stake (Fricker 1995). One ought, in other words, to be epistemically responsible in one's trusting (see Frost-Arnold 2020).

Such an epistemology is also open to criticisms, however. For example, it suggests that rational trust will always be partial rather than complete, given that the rational trustor is open to evidence that contradicts their trust on this theory, while someone who trusts completely in someone else lacks such openness. The theory also implies that the reasons for trusting well (i.e. in a justified way) are accessible to the trustor, at some point or another, which may be false. Some reasons for trust may be too "cunning" for this to be the case. Here, I have in mind the reason for trusting discussed by Philip Pettit (1995): that trust signals to people they are being held in esteem, which is something they will want to maintain; they will honor the trust because they are naturally "esteem-seeking." Consciously having this as a reason for trusting—of using people's need for esteem to get what you want from them—is incompatible with actually trusting (Wanderer and Townsend 2013, 9), if trust is motives-based and the required motive is something other than self-interest.

Others say that reasons for trust are usually too numerous and varied to be open to the conscious consideration of the trustor (e.g. Baier 1986). There can be very subtle reasons to trust or distrust someone—for example, reasons that have to do with body language, with systematic yet veiled forms of oppression, or with a complicated history of trusting others about which one couldn't easily generalize. Factors like these can influence trustors without them knowing it, sometimes making their trust irrational (e.g. because it is informed by oppressive biases), and other times making it rational.

The concern about there being complex reasons for trusting explain why some philosophers defend externalist epistemologies of trust. Some do so explicitly (e.g. McLeod 2002). They argue for reliabilist

theories that make trust rationally justified if and only if it is formed and sustained by reliable processes (i.e. "processes that tend to produce accurate representations of the world," such as drawing on expertise one has rather than simply guessing; Goldman 1992, 113; Goldman and Beddor 2016). Others gesture towards externalism (Webb 1993; Baier 1986), as Baier does with what she calls "a moral test for trust." The test is that "knowledge of what the other party is relying on for the continuance of the trust relationship would … itself destabilize the relation" (1986, 255). The other party might be relying on concealment of untrustworthiness or on a threat advantage, in which case the trust would probably fail the test. Because Baier's test focuses on the causal basis for trust, or for what maintains the trust relation, it is externalist. Also, because the trustor often cannot gather the information needed for the test without ceasing to trust the other person (Baier 1986, 260), the test cannot be internalist.

Although an externalist theory of trust deals well with some of the worries one might have with an internalist theory, it has problems of its own. One of the most serious problems is the absence of any requirement that trustors themselves have good reasons for trusting, especially when their trust makes them seriously vulnerable. Again, it appears that common sense dictates the opposite: that sometimes as trustors, we ought to be able to back up our decisions about when to trust. The same is true about our distrust presumably: that sometimes we ought to be able to defend it. Assuming externalists mean for their epistemology to apply to distrust and not just to trust, their theory violates this bit of common sense as well. Externalism about distrust also seems incompatible with a strategy that some philosophers recommend for dealing with biased distrust. The strategy is to develop what they call "corrective trust" (e.g. Scheman 2020) or "humble trust" (D'Cruz 2019), which demands a humble skepticism toward distrust that aligns with oppressive stereotypes and efforts at correcting the influence of these stereotypes (see also Fricker 2007). The concern about an externalist epistemology is that it does not encourage this sort of mental work, since it does not require that we reflect on our reasons for distrusting or trusting.

There are alternatives to the kinds of internalist and externalist theories I have discussed, especially within the literature on testimony.[12] For example, Paul Faulkner develops an "assurance theory" of testimony that interprets speaker trustworthiness in terms of trust-responsiveness. Recall that on a trust-responsiveness theory of trust, being trusted gives people the reason to be trustworthy that someone is counting on them. They are trustworthy if they are appropriately responsive to this reason, which, in the case of offering testimony, involves giving one's assurance that one is telling the truth (Adler 2017). Faulkner uses the trust-responsiveness account of trust, along with a view of trust as an affective attitude (see Trust and the Will), to show "how trust can ground reasonable testimonial uptake" (Faulkner and Simpson 2017, 6; Faulkner 2011 and 2020). "He proposes that A affectively trust S if and only if A depends on S Φ-ing, and expects his dependence on S to motivate S to Φ—for A's dependence on S to be the reason for which S Φs…. As a result, affective trust is a bootstrapping attitude: I can choose to trust someone affectively and my doing so creates the reasons which justify the attitude" (Faulkner and Simpson 2017, 6).

Most likely, A (the trustor) is aware of the reasons that justify his trust or could be aware of them, making this theory an internalist one. The reasons are also normative and non-evidentiary (Faulkner 2020); they concern what S ought to do because of A's dependence, not what S will do based on evidence that A might gather about S. This view doesn't require that A have evidentiary reasons, and so it is importantly different than the internalist epistemology discussed above. It is also subject as a result, however, to the criticisms made of externalist theories that they don't require the kind of scrutiny of our trusting attitudes that we tend to expect and that we ought to expect in societies where some people are stereotyped as more trusting than others.

Presumably to avoid having to defend any particular epistemology of trust, some philosophers provide just a list of common justifiers for it (i.e. "facts or states of affairs that determine the justification status of [trust]"; Goldman 1999, 274), which someone could take into account in deciding when to trust (Govier

1998, Jones 1996). Included on the lists are such factors as the social role of the trustee, the domain in which the trust occurs, an "agent-specific" factor that concerns how good a trustor the agent tends to be (Jones 1996, 21), and the social or political climate in which the trust occurs. Philosophers have tended to emphasize this last factor as a justification condition for trust, and so let me elaborate on it briefly.

2.3 Social and political climate

We have seen that although trust is paradigmatically a relation that holds between two individuals, forces larger than those individuals inevitably shape their trust and distrust in one another. Social or political climate contributes to how (un)trustworthy people tend to be and therefore to whether trust and distrust are justified. A climate of virtue, for example, is one in which trustworthiness tends to be pervasive, assuming that virtues other than trustworthiness tend to enhance it (Baier 2004).[13] A climate of oppression is one in which untrustworthiness is prevalent, especially between people who are privileged and those who are oppressed (Baier 1986, 259; Potter 2002, 24; D'Cruz 2019). "Social trust," as some call it, is low in these circumstances (Govier 1997, Welch 2013).

Social or political climate has a significant influence on the default stance we ought to take toward people's trustworthiness (see e.g. Walker 2006). We need such a stance—assuming some rational reflection is required for trusting well—because we can't always stop to reflect carefully on when to trust. Some philosophers say that the correct stance is trust and do so without referring to the social or political climate; Tony Coady takes this sort of position, for example, on our stance toward others' testimony (Coady 1992). Others disagree that the correct stance could be so universal and claim instead that it is relative to climate, as well as to other factors such as domain (Jones 1999).

Our trust or distrust may be prima facie justified if we have the correct default stance, although most philosophers assume that it could only be fully justified (in a truth- or end-directed way) by reasons that are internal to us (evidentiary or non-evidentiary reasons) or by the causal processes that created the attitude in the first place. Whichever epistemology of trust we choose, it ought to be sensitive to the tension that exists between trusting somebody and rationally reflecting on the grounds for that trust. It would be odd, to say the least, if what made an attitude justified destroyed that very attitude. At the same time, our epistemology ought to cohere as much as possible with common sense, which dictates that we should inspect rather than have pure faith in whatever makes us seriously vulnerable to other people, which trust can most definitely do.

3. The Value of Trust

Someone who asks, "When is trust warranted?" might be interested in knowing what the point of trust is. In other words, what value does it have? Although the value it has for particular people will depend on their circumstances, the value it could have for anyone will depend on why trust is valuable, generally speaking. Trust can have enormous instrumental value and may also have some intrinsic value. In discussing its instrumental value, I'll refer to the "goods of trust," which can benefit the trustor, the trustee, or society in general. They are therefore social as well as individual goods. What is more, they tend to accompany justified trust, rather than any old trust.[14] Like the other sections of this entry, this one will focus predominantly on trust, although I will also mention recent work on the value of distrust.

Consider first the possibility that trust has intrinsic value. If trust produced no goods independent of it, would there be any point in trusting? One might say "yes," on the grounds that trust is (or can be; O'Neil 2012, 311) a sign of respect for others. (Similarly, distrust is a sign of disrespect; D'Cruz 2019.) If true, this fact about trust would make it intrinsically worthwhile, at least so long as the trust is justified.

Presumably, if it was unjustified, then the respect would be misplaced and the intrinsic value would be lost. But I'm speculating here, since philosophers have said comparatively little about trust being worthwhile in itself as opposed to worthwhile because of what it produces, or because of what accompanies it. I'll say no more about the former and focus on the latter, in particular on the goods of trust.

Turning first to the instrumental value of trust to the trustor, some argue that trusting vastly increases our opportunities for cooperating with others and for benefiting from that cooperation, although of course we would only benefit if people we trusted cooperated as well (Gambetta 1988; Hardin 2002; Dimock 2020). Trust enhances cooperation, while perhaps not being necessary for it (Cook et al. 2005, Skyrms 2008). Because trust removes the incentive to check up on other people, it makes cooperation with trust less complicated than cooperation without it (Luhmann 1979).

Trust can make cooperation possible, rather than simply easier, if trust is essential to promising. Daniel Friedrich and Nicholas Southwood defend what they call the "Trust View" of promissory obligation (2011), according to which "making a promise involves inviting another individual to trust one to do something" (Friedrich and Southwood 277). If this view is correct, then cooperation through promising is impossible without trust. Cooperation of this sort will also not be fruitful unless the trust is justified.

Trusting provides us with goods beyond those that come with cooperation, although again, for these goods to materialize, the trust must be justified. Sometimes, trust involves little or no cooperation, so that the trustor is completely dependent on the trustee but the reverse is not true. Examples are the trust of young children in their parents and the trust of severely ill or disabled people in their care providers. Trust is particularly important for these people because they tend to be powerless to exercise their rights or to enforce any kind of contract. The trust they place in their care providers also contributes to them being vulnerable, and so it is essential that they can trust these people (i.e. that their trust is justified). The goods at stake for them are all the goods involved in having a good or decent life.

Among the specific goods that philosophers associate with trusting are meaningful relationships or attachments (rather than simply cooperative relationships that further individual self-interests; Harding 2011, Kirton forthcoming) as well as knowledge and autonomy.[15] To expand, trust allows for the kinds of secure attachments that some developmental psychologists ("attachment" theorists) believe are crucial to our well-being and to our ability to be trusting of others (Bowlby 1969, 1973, 1980; Ainsworth 1969; see Kirton forthcoming; Wonderly 2016). Particularly important here are parent-child relationships (McLeod et al. 2019).

Trust is also crucial for knowledge, given that scientific knowledge (Hardwig 1991), moral knowledge (Jones 1999), and almost all knowledge in fact (Webb 1993) depends for its acquisition on trust in the testimony of others. The basic argument for the need to trust what others say is that no one person has the time, intellect, and experience necessary to independently learn facts about the world that many of us do know. Examples include the scientific fact that the earth is round, the moral fact that the oppression of people from social groups different from our own can be severe (Jones 1999), and the mundane fact that we were born on such-in-such a day (Webb 1993, 261). Of course, trusting the people who testify to these facts could only generate knowledge if the trust was justified. If we were told our date of birth by people who were determined oddly to deceive us about when we were born, then we would not know when we were born.

Autonomy is another good that flows from trust insofar as people acquire or exercise autonomy only in social environments where they can trust people (or institutions, etc.) to support their autonomy. Feminists in particular tend to conceive of autonomy this way—that is, as a relational property (Mackenzie and Stoljar 2000). Many feminists emphasize that oppressive social environments can inhibit autonomy, and

some say explicitly that conditions necessary for autonomy (e.g. adequate options, knowledge relevant to one's decisions) exist only with the help of people or institutions that are trustworthy (e.g. Oshana 2014, McLeod and Ryman 2020). Justified trust in others to ensure that these conditions exist is essential for our autonomy, if autonomy is indeed relational.[16]

Goods of trust that are instrumental to the well-being of the trustee also do not materialize unless the trust is justified. Trust can improve the self-respect and moral maturity of this person. Particularly if it involves reliance on a person's moral character, trust can engender self-respect in the trustee (i.e. through them internalizing the respect signaled by that trust). Being trusted can allow us to be more respectful not only toward ourselves but also toward others, thus enhancing our moral maturity. The explicit goal of therapeutic trust is precisely to bring about this end. I suggested above (The Epistemology of Trust) that therapeutic trust can be justified in a truth-directed way over time, provided that the trust has its intended effect of making the trustee more trustworthy (McGeer 2008; Baker 1987, 12). Clearly, for therapeutic trust to benefit the trustee, it would have to be justified in this way, meaning that the therapy would normally have to work.

Finally, there are social goods of trust that are linked with the individual goods of cooperation and moral maturity. The former include the practice of morality, the very existence of society perhaps, as well as strong social networks. Morality itself is a cooperative activity, which can only get off the ground if people can trust one another to try, at least, to be moral. For this reason, among others, Baier claims that trust is "the very basis of morality" (2004, 180). It could also be the very basis of society, insofar as trust in fellow citizens to honor social contracts makes those contracts possible.

A weaker claim is that trust makes society better or more livable. Some argue that trust is a form of "social capital," meaning roughly that it enables "people to work together for common purposes in groups and organizations" (Fukuyama 1995, 10; quoted in Hardin 2002, 83). As a result, "high-trust" societies have stronger economies and stronger social networks in general than "low-trust" societies (Fukuyama 1995; Inglehart 1999). Of course, this fact about high-trust societies could only be true if, on the whole, the trust within them was justified—that is, if trustees tended not to "defect" and destroy chances for cooperating in the future.

The literature on distrust suggests that there are goods associated with it as well. For example, there is the social good discussed by Krishnamurthy of "securing democracy by protecting political minorities from tyranny" (2015, 392). Distrust as she understands it (a confident belief that others will not act justly) plays this positive role when it is justified, which is roughly when the threat of tyranny or unjust action is real. Distrust in general is valuable when it is justified—for the distrustors at least, who protect themselves from harm. By contrast, the people distrusted tend to experience negative effects on their reputation or self-respect (D'Cruz 2019).

Both trust and distrust are therefore valuable particularly when they are justified. The value of justified trust must be very high if without it, we can't have morality or society and can't be morally mature, autonomous, knowledgeable, or invested with opportunities for collaborating with others. Justified distrust is also essential, for members of minority groups especially. Conversely, trust or distrust that is unjustified can be seriously problematic. Unjustified trust, for example, can leave us open to abuse, terror, and deception.

4. Trust and the Will

Trust may not be warranted (i.e. plausible) because the agent has lost the ability to trust or simply cannot bring themselves to trust. People can lose trust in almost everyone or everything as a result of trauma

(Herman 1991). The trauma of rape, for example, can profoundly reduce one's sense that the world is a safe place with caring people in it (Brison 2002). By contrast, people can lose trust just in particular people or institutions or have never trusted them before. They or others might want them to become more trusting. But how could that happen? How can trust be restored or generated?

The process of building trust is often slow and difficult (Uslaner 1999; Baier 1986; Lahno 2020), and that is true, in part, because of the kind of mental attitude trust is. Many argue that it is not the sort of attitude we can simply will ourselves to have when it seems that the evidence in favour of it has vanished or never appeared to begin with. At the same, it is possible to cultivate trust.[17] I focus here on these issues, including what kind of mental attitude trust is (e.g. a belief or an emotion). I also discuss briefly what kind of mental attitude distrust is. Like trust, distrust is an attitude that people may wish to cultivate, particularly when they are too trusting.

Consider first why one would think that trust can't be willed. Baier questions whether people are able "to trust simply because of encouragement to trust" (1986, 244; my emphasis). She writes, "'Trust me!' is for most of us an invitation which we cannot accept at will—either we do already trust the one who says it, in which case it serves at best as reassurance, or it is properly responded to with, 'Why should and how can I, until I have cause to?'" (my emphasis; 1986, 244). Baier is not a voluntarist about trust, just as most people are not voluntarists about belief. In other words, she thinks that we can't simply decide to trust for purely motivational rather than epistemic reasons (i.e., merely because we want to, rather than because we have reason to think that the other person is or could be trustworthy; Mills 1998). That many people feel compelled to say, "I wish I could trust you," suggests that Baier's view is correct; wishing or wanting is not enough. But Holton interprets Baier's view differently. According to him, Baier's point is that we can never decide to trust, not that we can never decide to trust for motivational purposes (1994). This interpretation ignores, however, the attention that Baier gives to situations in which all we have is encouragement (trusting "simply because of encouragement"). The "cause" she refers to ("Why should and how can I, until I have cause to [trust]?"; 244) is an epistemic cause. Once we have one of those, we can presumably decide whether to trust on the basis of it.[18] But we cannot decide to trust simply because we want to, according to Baier.

If trust resembles belief in being non-voluntary, then perhaps trust itself is a belief. Is that right? Many philosophers claim that it is (e.g. Hieronymi 2008; McMyler 2011; Keren 2014), while others disagree (e.g. Jones 1996; Faulkner 2007; D'Cruz 2019). The former contend that trust is a belief that the trustee is trustworthy, at least in the thin sense that the trustee will do what he is trusted to do (Keren 2020). Various reasons exist in favour of such theories—doxastic ones (see Keren 2020)—including the reason that they make it impossible to trust a person while holding the belief that this person is not trustworthy, even in the thin sense. Most of us accept this impossibility and would want any theory of trust to explain it. A doxastic account does so by saying that we can't believe a contradiction (not knowingly anyway; Keren 2020, 113).

Those who say that trust is not a belief claim that it is possible to trust without believing the trustee is trustworthy.[19] Holton gives the nice example of trusting a friend to be sincere without believing that the friend will be sincere (1994, 75). Arguably, if one already believed that to be the case, then one would have no need to trust the friend. It is also possible to believe that someone is trustworthy without trusting that person, which suggests that trust could not just be a belief in someone's trustworthiness (McLeod 2002, 85). I might think that a particular person is trustworthy without trusting them because I have no cause to do so. I might even distrust them despite believing that they are trustworthy (Jones 1996, 2013). As Jones explains, distrust can be recalcitrant in parting "company with belief" (D'Cruz 2019, 940; citing Jones 2013), a fact which makes trouble for doxastic accounts not just of trust but of distrust as well (e.g. Krishnamurthy 2015). The latter must explain how distrust could be a belief that someone is

untrustworthy that could exist alongside the belief that the person is trustworthy.

Among the alternatives to doxasticism are theories stating that trust is an emotion, a kind of stance (i.e., the participant stance; Holton 1994), or a disposition (Kappel 2014; cited in Keren 2020). The most commonly held alternative is the first one: that trust is an emotion. Reasons in favour of this view include the fact that trust resembles an emotion in having characteristics that are unique to emotions, at least according to an influential account of them (de Sousa 1987; Calhoun 1984; Rorty 1980; Lahno 2001, 2020). For example, emotions narrow our perception to "fields of evidence" that lend support to the emotions themselves (Jones 1996, 11). When we are in the grip of an emotion, we therefore tend to see facts that affirm its existence and ignore those that negate it. To illustrate, if I am really angry at my mother, then I tend to focus on things that justify my anger while ignoring or refusing to see things that make it unjustified. I can only see those other things once my anger subsides. Similarly with trust: if I genuinely trust my mother, my attention falls on those aspects of her that justify my trust and is averted from evidence that suggests she is untrustworthy (Baker 1987). The same sort of thing happens with distrust (Jones 2019). Jones refers to this phenomenon as "affective looping," which, in her words, occurs when "a prior emotional state provides grounds for its own continuance" (2019, 956). She also insists that only affective-attitude accounts of trust and distrust can adequately explain it (2019).

There may be a kind of doxastic theory, however, that can account for the affective looping of trust, if not of distrust. Arnon Keren, whose work focuses specifically on trust, defends what he calls an "impurely doxastic" theory. He describes trust as believing in someone's trustworthiness *and* responding to reasons ("preemptive" ones) against taking precautions that this person will not be trustworthy (Keren 2020, 2014). Reasons for trust are themselves reasons of this sort, according to Keren; they oppose actions like those of carefully monitoring the behavior of the trustee or weighing the available evidence that this person is trustworthy. The trustor's response to these preemptive reasons would explain why this person is resistant (or at least not attune) to counter evidence to their trust (Keren 2014, 2020).

Deciding in favour of an affective-attitude theory or a purely or impurely doxastic one is important for understanding features of trust like affective looping. Yet it may have little bearing on whether or how trust can be cultivated. For, regardless of whether trust is a belief or an emotion, presumably we can cultivate it by purposefully placing ourselves in a position that allows us to focus on evidence of people's trustworthiness. The goal here could be self-improvement: that is, becoming more trusting, hopefully in a good way, so that we can reap the benefits of justified trust. Alternatively, we might be striving for the improvement of others: making them more trustworthy by trusting them therapeutically. Alternatively still, we could be engaging in "corrective trust." (See the above discussions of therapeutic and corrective trust.)

This section has centered on how to develop trust and how to account for facts about it such as the blinkered vision of the trustor. I have mentioned similar facts about distrust: those that concern what kind of mental attitude it is. Theorizing about whether trust and distrust are beliefs or emotions or what have you allows us to appreciate why they have certain features and also how to build these attitudes. The process for building them, which may be similar regardless of whether they are beliefs or emotions, will be relevant to people who don't trust enough or who trust too much.

5. Conclusion

This entry has examined an important practical question about trust: "When is trust warranted?" Also woven into the discussion has been some consideration of when distrust is warranted. Centerstage has been given to trust, however, because philosophers have debated it much more than distrust.

Different answers to the question of when trust is warranted give rise to different philosophical puzzles. For example, in response, one could appeal to the nature of trust and trustworthiness and consider whether the conditions are ripe for them (e.g. for the proposed trustor to rely on the trustee's competence). But one would first have to settle the difficult issue of what trust and trustworthiness are, and more specifically, how they differ from mere reliance and reliability, assuming there are these differences.

Alternatively, in deciding whether trust is warranted, one could consider whether it would be rationally justified or valuable. One would consider these things simultaneously when rational justification is understood in an end-directed way, making it dependent on trust's instrumental value. With respect to rational justification alone, puzzles arise when trying to sort out whether reasons for trust must be internal to trustors or could be external to them. In other words, is trust's epistemology internalist or externalist? Because good arguments exist on both sides, it's not clear how trust is rationally justified. Neither is it entirely clear what sort of value trust can have, given the nature of it. For example, trust may or may not have intrinsic moral value depending on whether it signals respect for others.

Lastly, one might focus on the fact that trust cannot be warranted when it is impossible, which is the case when the agent does not already trust and cannot simply will themselves to do so. While trust is arguably not the sort of attitude that one can just will oneself to have, trust can be cultivated. The exact manner or extent to which it can be cultivated, however, may depend again on what sort of attitude it is.

Since one can respond to the question, "When is trust warranted?" by referring to each of the above dimensions of trust, a complete philosophical answer to this question is complex. The same is true about the question of when to distrust, because the same dimensions (the epistemology of distrust, its value, etc.) are relevant to it. Complete answers to these broad questions about trust and distrust would be philosophically exciting and also socially important. They would be exciting both because of their complexity and because they would draw on a number of different philosophical areas, including epistemology, philosophy of mind, and value theory. The answers would be important because trust and distrust that are warranted contribute to the foundation of a good society, where people thrive through healthy cooperation with others, become morally mature human beings, and are not subject to social ills like tyranny or oppression.

Bibliography

Adler, J., 2017. "Epistemological Problems of Testimony," *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2017/entries/testimony-episprob/>.

Ainsworth, M. D. S., 1969. "Object Relations, Dependency, and Attachment: A Theoretical Review of the Infant-Mother Relationship," *Child Development*, 40(4): 969–1025.

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

——, 1991. "Trust and Its Vulnerabilities" and "Sustaining Trust," *Tanner Lectures on Human Values* (Volume 13), Salt Lake City: University of Utah Press.

——, 2004. "Demoralization, Trust, and the Virtues," in *Setting the Moral Compass: Essays by Women Philosophers*, C. Calhoun (ed.), New York, NY: Oxford University Press, pp. 176–189.

Baker, J., 1987. "Trust and Rationality," *Pacific Philosophical Quarterly*, 68: 1–13.

Becker. L. C., 1996. "Trust as Noncognitive Security about Motives," *Ethics*, 107(1): 43–61.

Blackburn, S., 1998. *Ruling Passion: A Theory of Practical Reasoning*, Oxford: Clarendon Press.

Bowlby, J., 1969. *Attachment and Loss* (Volume I: Attachment), London: The Hogarth Press and the Institute of Psycho-Analysis.

——, 1973. *Attachment and Loss* (Volume II: Separation, Anxiety and Anger), London: The Hogarth Press and the Institute of Psycho-Analysis.

––––, 1980. *Attachment and Loss* (Volume III: Loss, Sadness and Depression), London: The Hogarth Press and the Institute of Psycho-Analysis.

Brison, S., 2002. *Aftermath: Violence and the Remaking of the Self*, Princeton NJ: Princeton University Press.

Budnik, C., 2018. "Trust, Reliance, and Democracy, " *International Journal of Philosophical Studies*, 26: 221–239.

Calhoun, C. (ed.), 1984. "Cognitive Emotions?" in *What is an Emotion?*, C. Calhoun and R. C. Solomon (eds.), New York: Oxford University Press, pp. 327–342.

Coady, C.A.J., 1992. *Testimony: A Philosophical Study*, Oxford: Clarendon Press.

Coeckelbergh, M., 2012. "Can We Trust Robots?" *Ethics and Information Technology*, 14: 53–60.

Cogley, Z., 2012. "Trust and the Trickster Problem," *Analytic Philosophy*, 53(1): 30–47.

Cohen, M.A., and J. Dienhart, 2013. "Moral and Amoral Conceptions of Trust, with an Application to Organizational Ethics," *Journal of Business Ethics*, 112: 1–13.

Cook, K. R., Hardin, R. and M. Levi, 2005. *Cooperation Without Trust?*, New York: Russell Sage Foundation.

Dasgupta, P., 1988. "Trust as a Commodity," in Gambetta (ed.), pp. 49–72.

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

––––, 2011. "Altogether now: A virtue-theoretic approach to pluralism in feminist epistemology," in *Feminist Epistemology and Philosophy of Science*, H. Grasswick (ed.), New York: Springer, pp. 45–67.

D'Cruz, J., 2019. "Humble Trust," *Philosophical Studies*, 176: 933–953.

––––, 2020. "Trust and Distrust," in J. Simon (ed.), pp. 41–51.

de Sousa, R., 1987. *The Rationality of Emotion*, Cambridge: MIT Press.

Dimock, S., 2020. "Trust and Cooperation," in J. Simon (ed.), pp. 160–174.

Domenicucci, J. and R. Holton, 2017. "Trust as a Two-Place Relation," in P. Faulkner and T. Simpson (eds.), pp. 150–161.

Faulkner, P., 2007. "A Genealogy of Trust," *Episteme*, 4(3): 305–321.

––––, 2011. *Knowledge on Trust*, Oxford: Oxford University Press.

––––, 2015. "The Attitude of Trust is Basic," *Analysis*, 75(3): 424–429.

––––, 2017. "The Problem of Trust, " in P. Faulkner and T. Simpson (eds.), pp. 110–129.

––––, 2020. "Trust and Testimony," in J. Simon (ed.), pp. 329–340.

Faulkner, P. and T. Simpson (ed.), 2017. *The Philosophy of Trust*, Oxford: Oxford University Press.

Foley, R., 2001. *Intellectual Trust in Oneself and Others*, Cambridge: Cambridge University Press.

––––, 2005. "Universal Intellectual Trust," *Episteme*, 2(1): 5–11.

Fox-Decent, E., 2005. "The Fiduciary Nature of State Legal Authority," *Queen's Law Journal*, 31: 259–310.

Fricker, M., 2007. *Epistemic Injustice: Power and the Ethics of Knowing*, Oxford: Oxford University Press.

Fricker, E., 1995. "Telling and Trusting: Reductionism and Anti-Reductionism in the Epistemology of Testimony," *Mind*, 104 (414): 393–411.

Friedrich, D. and N. Southwood, 2011. "Promises and Trust," in *Promises and Agreements: Philosophical Essays*, H. Sheinman (ed.), Oxford: Oxford University Press, pp. 275–292.

Frost-Arnold, K., 2014. "Imposters, Tricksters, and Trustworthiness as an Epistemic Virtue," *Hypatia*, 29(4): 790–807.

––––, 2020. "Trust and Epistemic Responsibility," in J. Simon (ed.), pp. 64–75.

Fukuyama, F., 1995. *Trust: The Social Virtues and the Creation of Prosperity*, New York, NY: The Free Press.

Gambetta, D. (ed.), 1988. *Trust: Making and Breaking Cooperative Relations*, New York: Basil Blackwell.

––––, 1988. "Can We Trust Trust?" in D. Gambetta (ed.), pp. 213–237.

Goering, S., 2009. "Postnatal Reproductive Autonomy: Promoting Relational Autonomy and Self-Trust in New Parents," *Bioethics*, 23(1): 9–19.

Goldberg, S. C., 2020. "Trust and Reliance," in J. Simon (ed.), pp. 97–108.

Goldman, A. I., 1992. *Liaisons: Philosophy Meets the Cognitive and Social Sciences*, Cambridge, MA: MIT Press.

–––, 1999. "Internalism Exposed," *Journal of Philosophy*, 96(6): 271–293.

Goldman, A. and B. Beddor, 2016. "Reliabilist Epistemology," *Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/win2016/entries/reliabilism/>.

Govier, T., 1992a. "Trust, Distrust, and Feminist Theory," *Hypatia*, 7(1): 15–33.

–––, 1992b. "Distrust as a Practical Problem," *Journal of Social Philosophy*, 23(1): 52–63.

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

–––, 1997. *Social Trust and Human Communities*, Montreal and Kingston: McGill-Queen's University Press.

–––, 1998. *Dilemmas of Trust*, Montreal and Kingston: McGill-Queen's University Press.

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

––– (ed.), 2004, *Distrust*, New York: Russell Sage Foundation.

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

Hardwig, J., 1991. "The Role of Trust in Knowledge," *The Journal of Philosophy*, 88(12): 693–708.

Hawley, K., 2014. "Trust, Distrust and Commitment," *Noûs*, 48(1): 1–20.

–––, 2017. "Trustworthy Groups and Organizations," in P. Faulkner and T. Simpson (eds.), pp. 52–70.

–––, 2019. *How to Be Trustworthy*, Oxford: Oxford University Press.

Herman, J. L., 1991. *Trauma and Recovery*, New York: Basic Books.

Hertzberg, L., 1988. "On the Attitude of Trust," *Inquiry*, 31: 307–322.

Hieronymi, P., 2008. "The Reasons of Trust," *Australasian Journal of Philosophy*, 86(2): 213–236.

Hinchman, E. S., 2003. "Trust and Diachronic Agency," *Noûs*, 37(1): 25–51.

–––, 2017. "On the Risks of Resting Assured: An Assurance Theory of Trust," in P. Faulkner and T. Simpson (eds.), pp. 52–70.

Hollis, M., 1998. *Trust Within Reason*, Cambridge: Cambridge University Press.

Holton, R., 1994. "Deciding to Trust, Coming to Believe," *Australasian Journal of Philosophy*, 72(1): 63–76.

Horsburgh, H.J.N., 1960. "The Ethics of Trust," *Philosophical Quarterly*, 10: 343–354.

Inglehart, R., 1999. "Trust, Well-being and Democracy," in M. Warren (ed.), pp. 88–120.

Jacoby, B., 2011. *Trust and Betrayal: A Conceptual Analysis*, Ph.D. Dissertation, Philosophy Department, Macquarie University.

Jones, K., 1996. "Trust as an Affective Attitude," *Ethics*, 107: 4–25.

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

–––, 2004. "Trust and Terror," in *Moral Psychology: Feminist Ethics and Social Theory*, P. DesAutels and M. U. Walker (eds.), Lanham, Maryland: Rowman & Littlefield, pp. 3–18.

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

–––, 2012b. "The Politics of Intellectual Self-Trust," *Social Epistemology*, 26(2): 237–251.

–––, 2013. "Distrusting the Trustworthy," in *Reading Onora O'Neill*, D. Archard, M. Deveaux, N. Manson, and D. Weinstock (eds.), New York: Routledge.

–––, 2017. "But I Was Counting On You!" in P. Faulkner and T. Simpson (eds.), pp. 91–109.

–––, 2019. "Trust, Distrust, and Affective Looping," *Philosophical Studies*, 176: 955–968.

Kappel, K., 2014. "Believing on Trust," *Synthese*, 191: 2009–2028.

Keren, A., 2014. "Trust and Belief: A Preemptive Reasons Account," *Synthese*, 191: 2593–2615.

–––, 2020. "Trust and Belief," in J. Simon (ed.), pp. 109–120.

Kirton, A., Forthcoming. "Matters of Trust as Matters of Attachment Security," *International Journal of Philosophical Studies*,

Krishnamurthy, M., 2015. "White Tyranny and the Democratic Value of Distrust," *The Monist*, 98(4): 392–406.

Lagerspetz, O., 1998. *Trust: The Tacit Demand*, Dordrecht: Kluwer.

Lahno, B., 2001. "On the Emotional Character of Trust," *Ethical Theory and Moral Practice*, 4: 171–189.

–––, 2020. "Trust and Emotion," in J. Simon (ed.), pp. 147–159.

Lehrer, K., 1997. *Self-Trust: A Study of Reason, Knowledge, and Autonomy*, New York: Oxford University Press.

Luhmann, N., 1979. *Trust and Power*, Toronto: Wiley.

Mackenzie, C. and N. Stoljar (eds.), 2000. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, New York: Oxford University Press.

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

McGeer, V. and P. Pettit, 2017. "The Empowering Theory of Trust," in P. Faulkner and T. Simpson (eds.), pp. 15–35.

McLeod, C., 2002. *Self-Trust and Reproductive Autonomy*, Cambridge, MA: MIT Press.

–––, 2020. "Damage to Trust," in her *Conscience in Reproductive Health Care: Prioritizing Patient Interests*, Oxford: Oxford University Press, pp. 65–87.

McLeod, C. and E. Ryman, 2020. "Trust, Autonomy, and the Fiduciary Relationship," in *Fiduciaries and Trust: Ethics, Politics, Economics, and Law*, P. B. Miller and M. Harding (eds.), Cambridge: Cambridge University Press, pp. 74–86.

McLeod, C., L. Davies, N. Fice, L. Bruijns, E. Cichocki, H. Doguoglu, H. Stewart, A. Horn, J. Rekis, T. Filipovich, 2019. "Time to Attach: An Argument in Favour of EI Attachment Benefits," report created for Ontario's Adoptive Parents Association and the Adoption Council of Canada, available online.

McMyler, B., 2011. *Testimony, Trust, and Authority*, Oxford: Oxford University Press.

Mills, E., 1998. "The Unity of Justification," *Philosophy and Phenomenological Research*, 58(1): 27–51.

Mullin, A., 2005. "Trust, Social Norms, and Motherhood," *Journal of Social Philosophy*, 36(3): 316–330.

Nickel, P. J., 2007. "Trust and Obligation-Ascription," *Ethical Theory and Moral Practice*, 10: 309–319.

O'Neil, C., 2012. "Lying, Trust, and Gratitude," *Philosophy and Public Affairs*, 40(4): 301–333.

O'Neill, O., 2002. *Autonomy and Trust in Bioethics*, Cambridge: Cambridge University Press.

Oreskes, N., 2019. *Why Trust Science?* Princeton NJ: Princeton University Press.

Oshana, M., 2014. "Trust and Autonomous Agency," *Res Philosophica*, 91(3): 431–447.

Pellegrino, E. D., and D. C. Thomasma, 1993. *The Virtues in Medical Practice*, New York: Oxford University Press.

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

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

–––, 2013. "Narrative Selves, Relations of Trust, and Bipolar Disorder," *Philosophy, Psychiatry, & Psychology*, 20(1): 57–65.

–––, 2020. "Interpersonal Trust," in J. Simon (ed.), pp. 243–255.

Rorty, A. O., 1980. "Explaining Emotions," in *Explaining Emotions*, A. O. Rorty (ed.), Berkeley: University of California Press.

Scheman, N., 2020. "Trust and Trustworthiness," in J. Simon (ed.), pp. 28–40.

Skyrms, B., 2008. "Trust, Risk, and the Social Contract," *Synthese*, 160: 21–25.

Simon, J. (ed.), 2020. *The Routledge Handbook of Trust and Philosophy*, New York: Routledge.

Simpson, T. W., 2012. "What is Trust?" *Pacific Philosophical Quarterly*, 93: 550–569.

Smith, M. N., 2008. "Terrorism, Shared Rules and Trust," *The Journal of Political Philosophy*, 16(2): 201–219.

Strawson, P. F., 1974. "Freedom and Resentment," in his *Freedom and Resentment*, London: Methuen, pp. 1–25.

Sullins, J. P., 2020. "Trust in Robots," in J. Simon (ed.), pp. 313–325.

Tallant, J., 2017. "Commitment in Cases of Trust and Distrust," *Thought*, 6: 261–267.

Townley, C., and J. L. Garfield, 2013. "Public Trust," in *Trust: Analytic and Applied Perspectives*, P. Makela and C. Townley (eds.), Amsterdam: Rodopi Press, pp. 95–107.

Ullmann-Margalit, E., 2017. "Trust, Distrust, and In Between," in *Normal Rationality: Decisions and Social Order*, E. Ullmann-Margalit, A. Margalit, and C. R. Sunstein (eds.), Oxford: Oxford University Press, pp. 184–202.

Uslaner, E. M., 1999. "Democracy and Social Capital," in M. Warren (ed.), pp. 121–150.

Walker, M. U., 2006. *Moral Repair: Reconstructing Moral Relations After Wrongdoing*, Cambridge: Cambridge University Press.

Wanderer, J., and L. Townsend, 2013. "Is it Rational to Trust?" *Philosophy Compass*, 8(1): 1–14.

Warren, M. (ed.), 1999. *Democracy and Trust*, Cambridge: Cambridge University Press.

Webb, M. O., 1992. "The Epistemology of Trust and the Politics of Suspicion," *Pacific Philosophical Quarterly*, 73: 390–400.

–––, 1993. "Why I Know About as Much as You: A Reply to Hardwig," *The Journal of Philosophy*, 90(5): 260–270.

Welch, S., 2013. "Transparent Trust and Oppression," *Critical Review of International Social and Political Philosophy*, 16(1): 45–64.

Wonderly, M. L., 2016. "On Being Attached," *Philosophical Studies*, 173(1): 223–242.

Zaner, R., 1991. "The Phenomenon of Trust in the Patient-Physician Relationship," in *Ethics, Trust, and the Professions: Philosophical and Cultural Aspects*, E. D. Pellegrino, R.M. Veatch, and J.P. Langan (eds.), Washington, DC: Georgetown University Press, pp. 45–67.

Academic Tools

📝 How to cite this entry.

📝 Preview the PDF version of this entry at the Friends of the SEP Society.

📖 Look up this entry topic at the Internet Philosophy Ontology Project (InPhO).

PP Enhanced bibliography for this entry at PhilPapers, with links to its database.

Other Internet Resources

- 'A Question of Trust,' BBC Reith Lectures.
- Russell Sage Foundation Program on Trust.
- Building Trust, *OECD Observer*.
- 'Trust and Mistrust,' Philosophy Talk.
- 'On Trust and Philosophy,' OpenLearn.
- 'The Philosophy of Trust: Key Findings,' The Trust Project at Northwestern University.

Related Entries

Aristotle, General Topics: ethics | autonomy: in moral and political philosophy | autonomy: personal | emotion | epistemology: social | faith | feminist philosophy, interventions: social epistemology | feminist philosophy, topics: perspectives on autonomy | feminist philosophy, topics: perspectives on the self | friendship | justification, epistemic: internalist vs. externalist conceptions of | reliabilist epistemology | rights: of children | scientific knowledge: social dimensions of | testimony: epistemological problems of

Acknowledgments