

## TRUST AND BELIEF: A PREEMPTIVE-REASONS ACCOUNT

Arnon Keren

### **Abstract**

According to doxastic accounts of trust, trusting a person to  $\Phi$  involves, among other things, holding a belief about the trusted person: either the belief that the trusted person is trustworthy or the belief that she actually will  $\Phi$ . In recent years, several philosophers have argued against doxastic accounts of trust. They have claimed that the phenomenology of trust suggests that rather than such a belief, trust involves some kind of non-doxastic mental attitude towards the trusted person, or a non-doxastic disposition to rely upon her. This paper offers a new account of reasons for trust and employs the account to defend a doxastic account of trust. The paper argues that reasons for trust are preemptive reasons for action or belief. Thus the Razian concept of preemptive reasons, which arguably plays a key role in our understanding of relations of authority, is also central to our understanding of relations of trust. Furthermore, the paper argues that acceptance of a preemptive account of reasons for trust supports the adoption of a doxastic account of trust, for acceptance of such an account both neutralizes central objections to doxastic accounts of trust and provides independent reasons supporting a doxastic account.

Keywords: Trust, Belief, Reason, Testimony, Authority

We trust others to do various things. Tony trusts the babysitter to take good care of the kids. Tanya trusts her employee not to steal from the till. Todd trusts his wife to be faithful to him.

One thing we can trust another person to do is to tell us the truth: a speaker might tell us that  $p$ , and we may simply take her word for it. I shall call this form of trust 'speaker-trust', and beliefs formed in this way 'trust-based beliefs.'

As the case of speaker-trust suggests, trust can be a source of belief. But does trust itself involve a belief? What are we saying of Tony, when we say that he trusts the babysitter to take good care of the kids? According to doxastic accounts of trust, what we are ascribing to Tony is, among other things, a belief about the babysitter: either the belief that she is trustworthy, or that she will take good care of the kids.

While early philosophical treatments of trust often assumed a doxastic account (Baier 1986, Adler 1994), in recent years several philosophers have argued against such accounts, claiming that the phenomenology of trust suggests that trust does not always involve a belief. Instead, they have claimed, trust involves some other non-doxastic mental attitude—perhaps an affective attitude, or the adoption of the participant's stance toward the trustee—or simply a non-doxastic disposition to rely upon the trusted person (Jones 1996; McLeod 2011; McLeod 2002; Holton 1994; Faulkner 2007; Kappel, forthcoming).<sup>1</sup>

As noted by Jones, what we say about the nature of trust and of the mental attitude, if any, required by trust will make a difference to what we say about the conditions under which trust is rational (1996, 4–5). Moreover, as recent literature

---

<sup>1</sup> Non-doxastic accounts need not deny that trusting often involves a belief about the trusted person or even that it is sometimes in virtue of having such a belief that one can be described as trusting. For example, Frost-Arnold (forthcoming) suggests an account according to which  $A$  trusts  $B$  to  $\Phi$  if she either believes or accepts the proposition that  $B$  will  $\Phi$ . For our purposes, this is nevertheless a non-doxastic account, because it denies that trust entails belief.

suggests, the relation between these two questions also works in the other direction: what we say about what justifies trust may make a difference to what we say about the nature of trust. Thus, an important virtue which Jones ascribes to her non-doxastic account of trust is that it implies that the rationality of trust is not determined by what normally (always?) determines the rational status of a belief: namely evidence (1996). Indeed, some of the central arguments against doxastic accounts of trust are based on the claim that what justifies trust is very much unlike what justifies belief:<sup>2</sup> that there is a principled difference between the relations between belief and evidence and those between trust and evidence, or even a tension between believing on trust and believing on evidence.

In this paper I want to defend a doxastic account of trust, based on a claim about the nature of reasons for trust: I will argue that reasons for trust are preemptive reasons for action or belief. Thus the Razian concept of *preemptive reasons*, which arguably plays a key role in our understanding of relations of authority (Raz 1990a), is also central to our understanding of trust. This account of reasons for trust both neutralizes several objections to doxastic accounts of trust and provides independent reasons for the adoption of a doxastic account.<sup>3</sup> In particular, it supports the claim that one cannot trust a person unless one believes that she is trustworthy.<sup>4</sup>

---

<sup>2</sup> Even some supporters of doxastic accounts concede that the kind of trusting beliefs that are central to trust are not based upon evidence (Hieronymi 2008), and that accounting for some of trust's distinctive features requires severing the relation between trust and evidence. According to the account defended here, this is not the case.

<sup>3</sup> Both Keren (2007) and Zagzebski (2012) explore the idea that forms of epistemic trust involve responding to preemptive reasons for belief. However, they only address epistemic trust; neither suggests that preemptive reasons play a role in our understanding of trust in general nor that this notion can shed light on the debate between supporters of doxastic and non-doxastic accounts of trust. For further discussion of the difference between the view presented here and Zagzebski's, see footnote 15.

<sup>4</sup> Note that I make the claim that trust entails a belief, not the stronger claim that trust is a belief. Thus, the account defended here is doxastic, but not purely doxastic. Trust, I claim, requires not merely believing that the trustee is trustworthy but also responding to preemptive reasons (that one sees

In section 1, I start out by discussing some features of trust on which central objections to doxastic accounts are based. Section 2 turns to two challenges facing non-doxastic accounts when they are applied to the case of speaker-trust, arguing that doxastic accounts have an obvious advantage here. To show that doxastic accounts can maintain this advantage elsewhere, I argue that a preemptive-reasons account of reasons for trust can best explain some of the distinctive features of speaker-trust (sections 3 and 4) and trust in general (sections 5 and 6), including those features that appear to tell against doxastic accounts of trust. I then argue that the preemptive-reasons account of reasons for trust fits best within a doxastic account of trust (section 7). Finally, I discuss some implications of this for the rationality of trust and trust-based beliefs (section 8).

## **1. Trust, Reliance and Evidence**

Several philosophers have observed that the relation between trust and evidence is different from that between belief and evidence. Indeed, some have suggested that there is a tension, or even incompatibility, between believing or acting on trust and believing or acting on evidence:

[T]rust need not satisfy either a positive or a negative evidence condition: it need not be based on evidence and can demonstrate a wilful insensitivity to the evidence. Indeed there is a tension between acting on trust and acting on evidence that is illustrated in the idea that one does not actually trust someone to do something if one only believes they will do it when one has evidence that they will. (Faulkner 2007, 876)

Before examining this apparent tension, we should note some basic points about the notion of trust. As is standard in the literature, I shall think of trust as a three-place

---

oneself as having in virtue of what one believes about the trustee) against taking certain precautions. Moreover, my (partial) account of trust does not rule out the possibility that trust also involves having some non-doxastic attitude in addition to belief.

relation: person  $A$  trusts person  $B$  to  $\Phi$ .<sup>5</sup> And I assume that there is a difference between trusting  $B$  to  $\Phi$  and merely relying on her to  $\Phi$  (Baier 1986). In both cases, we work the supposition that she will  $\Phi$  into our plans (Holton 1994), thereby incurring the risk that she might fail to  $\Phi$ . Yet trust "is a special kind of reliance" (Holton 1994, 64), and we can rely on a person to  $\Phi$  without trusting her to do so (Baier 1986).

The distinction between trust and reliance is significant for the alleged tension between trust and evidence. We can surely rely on another person as we rely on evidence (as Kant's neighbors relied on his regular habits as an indication of the time). Similarly, we can rely on a person's testimony as we rely on any other piece of evidence (inferring, for instance, from the testimony of someone who is systematically confused that his testimony happens to be true). However, since we can rely without trusting, the fact that we can rely on  $B$  to  $\Phi$  while believing on the basis of evidence that she will  $\Phi$  does not entail that we can also trust  $B$  to  $\Phi$  if we only believe that she will  $\Phi$  when we have supporting evidence. Indeed, several features of trust (its resistance to counterevidence, the fact that it can be undermined by rational reflection and the phenomenon of therapeutic trust) seem to suggest that trust and belief differ in their relation to evidence. It is these differences that may suggest that I do not trust a person if I rely on her (or on her testimony) in the same way that I rely on any other piece of evidence (Hieronymi 2008, Faulkner 2007), and that have led several philosophers to reject doxastic accounts of trust.

---

<sup>5</sup> Trustworthiness can also be understood as having a three-place structure (Jones 2012): person  $B$  is trustworthy with respect to person  $A$  and action  $\Phi$  (or domain of interaction  $D$ ). Whether a three-place relation of trust is the most basic trusting relation is a question I will not address here; on this see Jones (2004).

While non-doxastic accounts differ from each-other in important respects, what is common to them all is the claim that what we ascribe to a person, when we say of him that he trusts  $B$  to  $\Phi$ , is not a belief about  $B$ : Trusting  $B$  to  $\Phi$  neither entails believing that  $B$  is trustworthy, nor that she will  $\Phi$ .<sup>6</sup> According to an affective-attitude account of trust, what we ascribe to Tony when we say that he trusts the babysitter to take good care of the kids is an affective attitude of optimism about the babysitter's goodwill and competence (Jones 1996): an attitude of optimism that, even though it tends to give rise to beliefs, is not to be cashed out in terms of belief, but in terms of an affective attitude. According to a participant-stance account (Holton 1994), what we are saying is that Tony relies on her to take good care of them while adopting the participant's stance toward her (which involves being prepared to feel certain reactive attitudes should he find out that his trust was betrayed). Neither reliance nor the adoption of the participant's stance, Holton insists, requires believing that the babysitter is trustworthy or that she will do what she is trusted to do. According to a dispositional account of trust, of the kind suggested by Kappel with respect to speaker-trust (forthcoming), to trust someone is simply to have a certain non-doxastic disposition to rely upon her (and in the case of speaker-trust, specifically, to accept what she says as true).<sup>7</sup>

---

<sup>6</sup> Jones, while committing herself to the claim that trust does not entail a belief (1996, 22), also presents a weaker claim—that "trust is not *primarily* a belief" (1996, 5; emphasis added). This seems to be consistent with a mixed doxastic account, according to which trust involves some non-doxastic attitude as a central component alongside a non-central doxastic component, where the centrality of the non-doxastic component is to be understood, at least partially, in terms of its role in explaining features of trust. Without arguing here against mixed accounts of trust, I do want to insist on the centrality of the belief component: as I shall argue, features of trust which Jones attempts to explain by appealing to non-doxastic components are better explained by appealing to the belief component and to the preemptive reasons which an agent has by virtue of her belief.

<sup>7</sup> Kappel does not presuppose that trust is a distinct or uniform type of attitude (forthcoming). Accordingly, unlike other accounts, his does not attempt to draw a line distinguishing trust from other attitudes. Nonetheless, a dispositional account of the kind he proposes may suggest a way of doing so.

Of the features of trust that have led philosophers to adopt such non-doxastic accounts, the first is trust's resistance to counter-evidence (Jones 1996, McLeod 2011, McLeod 2002, Faulkner 2007, Faulkner 2011, Baker 1987).<sup>8</sup> Thus, if I genuinely trust my friend to distribute money to the needy, I will tend to disbelieve accusations that she has embezzled that money. Moreover, I will not arrive at such disbelief by weighing evidence for her honesty against counter-evidence (Baker 1987, 3). This insensitivity to evidence supporting accusations against trusted friends appears different from that exhibited by confident beliefs that are immune from doubt because they are supported by an extensive body of evidence. For our disbelief in such accusations is often not based on extensive evidence against the veracity of the accusations. Both Jones (1996) and McLeod (2002) argue that in this respect, trust is similar to an emotion: having an emotion involves being attuned to certain information, along with a tendency to ignore other information that goes against the emotion.

The second such feature is that trust can be undermined by rational reflection. "Trust is a fragile plant, which may not endure inspection of its roots, even when they were, before the inspection, quite healthy" (Baier 1986, 260). Trust involves the acceptance of risk; an attempt to eliminate that risk by excessive reflection on relevant considerations tends to undermine trust (McLeod 2011; Faulkner 2007).<sup>9</sup> With belief, by contrast, if we have supporting evidence and reflect on it, this tends to strengthen, not undermine, the belief. If reflection *per se* undermines trust, does this not suggest that the attitude involved in trust is not a belief?

---

<sup>8</sup> Baker (1987), while noting that trust is resistant to evidence, does not argue for a non-doxastic account of trust. Instead, she suggests that this feature should make us reconsider the standards of rationality that apply to beliefs.

<sup>9</sup> Moreover, as the above quotation from Baier suggests, it can undermine trust even in cases where initially we had strong evidence supporting the trustee's trustworthiness

The third consideration against doxastic accounts emerges from what has been called 'therapeutic trust': cases where we seem to place our trust in someone in order to promote her trustworthiness, not because we have evidence for her trustworthiness (Holton 1994, McLeod 2002, Jones 1996, Faulkner 2007, Faulkner 2012, Frost-Arnold, forthcoming; Pettit 1995). Jones gives the example of parents who trust their teenage daughter to look after the house despite her past failures to fulfill similar responsibilities, hoping by their trust to eventually elicit more responsible and trustworthy behavior (2004, 5). Such cases seem to suggest that we can trust a person to  $\Phi$  without believing that she is trustworthy or that she will  $\Phi$ .<sup>10</sup>

## 2. The Case of Speaker-trust

While some features of trust seem to tell against doxastic accounts of trust, doxastic accounts also seem to have certain advantages over non-doxastic ones. Speaker-trust offers a useful testing ground, for while it has some of the features that might seem to tell against doxastic accounts of trust,<sup>11</sup> speaker-trust also involves forming a belief, and systematically so. If a speaker tells us that  $p$ , and we take her word for it, then we believe that  $p$ . This, I shall argue, provides doxastic accounts with an obvious advantage.

Some supporters of non-doxastic accounts of trust deny that trusting a speaker for the truth of what she says implies believing what she says. Faulkner (2007) suggests that trusting a speaker involves accepting her testimony, where accepting a proposition involves adopting a policy of treating it as true but need not involve

---

<sup>10</sup> Moreover, if the parents' trust in their daughter does not involve any belief in her trustworthiness, then such trust may appear to be rational: lack of evidence for the daughter's trustworthiness need not tell against the rationality of their trust.

<sup>11</sup> It is controversial whether speaker-trust has all of these features and, in particular, whether we can therapeutically trust a speaker for the truth. Compare Hieronymi (2008) and Faulkner (2007).



believing it. However, it is highly questionable whether such acceptance of testimony amounts to trusting the speaker. As Faulkner himself admits, speakers' expectations that we trust them would arguably not be satisfied by such treatment of their testimony (2007, 894).<sup>12</sup>

Most supporters of non-doxastic accounts, including Faulkner in his recent work (2011), therefore accept the systematic relation between trusting a speaker and believing what she says. They insist, however, that the explanation of this relationship is *not* the one offered by what Moran (2005) describes as evidentialist accounts of testimony. On such an account, in trusting a speaker, we believe that she is trustworthy—i.e. knowledgeable and honest—and therefore take her testimony that *p* as good evidence for *p*, and therefore believe that *p*. Instead, they claim that while trusting a speaker need not involve believing that she is trustworthy, the attitude or disposition involved in trust *results* in the formation of a belief that her testimony is true (Holton 1994; Faulkner 2011; Jones 1996).

However, this claim faces two main challenges. First, its supporters must explain why trusting a speaker to speak knowledgeably and sincerely, without believing that she will, should result in believing what she says. Moreover, if they wish to avoid saying, with Faulkner, that one can take a speaker's word for it without believing what she says, they must explain why trust should have such a result in each and every case. It won't do to suggest that the kind of non-doxastic mental state involved in trust has a tendency to give rise to beliefs in the way emotions tend to give

---

<sup>12</sup>Frost-Arnold (forthcoming), while agreeing that merely acting as if a speaker's testimony is true does not amount to trusting her, argues that this is no objection to accounts of trust suggesting that trusting speakers sometimes involves accepting their testimony without believing it, because merely acting as if a proposition is true does not amount to accepting it. However, this does not avoid the above objection. Since acceptance, unlike belief, is context-dependent, we can accept a speaker's testimony in one context while asserting, in other contexts, that we don't know whether her testimony is true. Surely such acceptance would not satisfy speakers' expectations that we trust them.

rise to beliefs (Jones 1996): For we can have an emotion, such as fear, without having the belief in the dangerousness of what we fear. Thus, without positing that trusting a person to  $\Phi$  invariably involves believing that she will  $\Phi$ , it is unclear how to explain this invariable relation between trusting a speaker and believing what she says.

The second main challenge is this: suppose that David believes that  $p$  because Susan told him so and he trusts her. And suppose that a non-doxastic account can explain why his trusting her results in his believing that  $p$ , without attributing to him the belief that she is trustworthy. Can this account also explain how David can see his own resulting belief as trust-based without the belief thereby being undermined? Arguably, we cannot hold onto a belief if we find that it is not supported by truth-conducive reasons and that what caused us to hold the belief has nothing to do with its truth. But we are generally able to hold on to trust-based beliefs while believing that we hold them because we trust the speaker. It would seem that to see his belief as supported by truth-conducive reasons, David must either believe that Susan is trustworthy or take his belief to be supported by independent evidence. But in the latter case he would not see his belief as based upon trust.

Doxastic accounts show an obvious advantage on both counts: they can explain the systematic relation between trusting the speaker and believing what she says and they can explain how David can see his trust-based belief as supported by truth-conducive reasons of the most paradigmatic kind, namely evidence: Because he believes that Susan is trustworthy, he sees her testimony itself as good evidence for the attested proposition.

The question, however, is whether a doxastic account can maintain its advantage, once we consider those features of trust discussed in section 1. To the extent that doxastic accounts employ an evidentialist explanation, as described above,

can they explain the difference between the phenomenology of speaker-trust and that of other evidence-based beliefs?

### **3. Reasons for Trust**

To allow us to consider this question, compare the following two cases. Consider, first, a speaker, Sandy, who addresses us and says: "take my word for it,  $p$ ." And suppose that we form the belief that  $p$ , and that we do so without weighing other relevant evidence available to us. It seems that Sandy cannot now turn to us and say: "You are correct. But you should have considered all the evidence available to you before reaching that conclusion." Criticizing us in this way appears incompatible with her expectation of us that we take her word for it. If she were to criticize us in this way, she would seem to be withdrawing her original appeal to us that we take her word for it.

Now compare this with the case of Evelyn, who intentionally provides us with very good evidence for  $p$ , expecting us to recognize her intention to provide us with such evidence. Suppose that here we again form the belief that  $p$ , without weighing other available evidence. If Evelyn's only intention was to provide us with very good evidence for  $p$ , then it seems that she can criticize us by saying: "You are correct. But you should have considered all the available evidence before reaching that conclusion." That is, criticizing us for not weighing other evidence is perfectly consistent with her expectation of us that we recognize that she has provided us with very good evidence. While such criticism does not appear to be compatible with a speaker's expectation of us that we take her word for it, it does appear to be compatible with a person's expectation that we recognize that she has provided us with good evidence.

How can we account for this difference within an account of speaker-trust suggesting that, in getting told and being believed, evidence is provided by speakers and accepted by audiences? If our account employs an evidentialist explanation for why trusting a speaker involves believing what she says, then it would suggest that in inviting us to trust her, Sandy, no less than Evelyn, expects us to recognize that she has provided us with good evidence. Why is it then that criticizing us for not weighing other evidence seems compatible with Evelyn's expectations, but not with Sandy's?

Without giving up the idea that trusting a speaker involves seeing her testimony as good evidence, we can explain the difference by saying that by inviting us to trust her about  $p$ , a speaker purports to provide us with both evidence for  $p$  and a preemptive reason for belief: that is, with a second-order reason for not forming our opinion regarding  $p$  on the basis of our own weighing of certain other evidence that may be available to us. In a way, the testimony of a trusted speaker is analogous to an authoritative command, in the following limited respect. An authoritative command gives us good reason to act, but not only that. It also gives us, or purports to give us, what Raz (1990a) has labeled a *preemptive reason* for action: a higher-order reason against acting for certain other reasons.<sup>13</sup> Similarly, a speaker inviting us to trust her purports to provide us with evidence, but not only that. She also purports to provide us with a higher-order reason against forming a belief on the basis of our own

---

<sup>13</sup> Raz coins the term "preemptive reason" to describe a reasons that "displaces" other reasons (1990a, 10; 1986, 42). At other places Raz describes them as reasons that "exclude" or "preempt" other reasons, and in some writings refers to them as "exclusionary reasons". The core idea of all these expressions is that such reasons are second-order reasons against acting on certain first-order reasons. It is in this sense that I shall use the term here.

Note however that the term "preemptive reasons" is sometimes used, for instance, by Darwall (2010), as interchangeable with what Raz calls "protected reasons". In this sense, a preemptive or protected reason is both a first-order reason to act in a certain way and a second-order exclusionary reason not to act on conflicting first-order reasons (Raz 1990b). To avoid any confusion, I shall use the term "protected reasons" to refer to this kind of combination of first- and second-order reasons. As I explain below (footnote 20), reasons for trust are also protected reasons.

weighing of certain other evidence.<sup>14</sup> It is because she purports to provide us with a reason against weighing other evidence that Sandy cannot criticize us for not weighing such evidence; if she nonetheless criticized us for that she would appear to be withdrawing her original appeal.

A speaker who invites you to trust her is inviting you not merely to rely upon her judgment but to rely on it, more specifically, *instead* of making your own judgment on the basis of all the evidence available to you. She presents herself as trustworthy, as honest and knowledgeable, and as having taken precautions against misleading you; and because she is thus trustworthy, you can believe her testimony without weighing all the evidence. In this sense, she purports to provide you with a preemptive reason for belief. I want to suggest that what distinguishes speaker-trust from mere reliance is not that in trusting speakers one does not see their testimony as good evidence for  $p$  but rather that the speaker's testimony is seen *not only* as good evidence but also as a preemptive reason for believing that  $p$ .

The obvious advantage of this explanation over that provided by non-evidentialist and non-doxastic accounts of speaker-trust is that it accounts for the distinctiveness of speaker-trust without severing the connection between reasons for trust-based beliefs and evidence. There is then no mystery in the fact that speaker-trust results in our forming a belief nor in the fact that we take our resulting beliefs to be epistemically rational and supported by truth-conducive reasons.

There is, however, yet another requirement which the preemptive-reasons account of speaker-trust must meet: that of explaining why, on this account, audiences who form a belief through speaker-trust see their belief as based upon *trust*. What, on this account, makes speaker-trust a genuine relation of trust?

---

<sup>14</sup> For a discussion of the concept of preemptive reasons for belief, see Zagzebski (2012).

I want to suggest that reasons for trust *in general* are preemptive reasons for belief or for action. The feature of speaker-trust noted here is an instance of a general feature of trust, one that distinguishes trust in general from mere reliance. To trust someone is, as Elster has suggested, "to lower one's guard, to refrain from taking precautions against an interaction partner, even when the other, because of opportunism or incompetence, could act in a way that might seem to justify precautions" (2007, 344). The fact that, in trusting others, we refrain from taking precautions means that reasons for trust are higher-order reasons against acting for certain other reasons—precautionary reasons—and hence that they are preemptive reasons.

In relying on others to form a belief, we incur the risk of forming a false belief. The main precaution which we can take against this risk involves weighing available evidence. But reasons for trust are reasons against taking precautions and, therefore, against weighing such evidence. Thus the above account of speaker-trust readily follows from a general account according to which reasons for trust are preemptive reasons. As I shall now argue this general account is better able than are other accounts of trust to explain features of speaker-trust and of trust. Indeed, it better explains some of the very features of trust that have attracted some to non-doxastic accounts of trust.<sup>15</sup>

---

<sup>15</sup> Both Zagzebski (2012) and McMyler (2011) propose accounts of epistemic trust, and both employ in this context ideas about the preemptive nature of beliefs on authority. Zagzebski explicitly adopts the notion of "preemption" derived from Raz (1986), and the notion plays a central role in her account of believing on the epistemic authority of others. However, unlike the account of trust suggested here, the notion of preemption does not figure in her understanding of trust *per se* and of epistemic-trust in particular. While on her account, epistemic authority has everything to do with preemption, trust *per se*, whether epistemic or non-epistemic, does not. Instead, preemption figures within her discussion of epistemic trust, because "among those we are committed to trusting are some whom we ought to treat as epistemic authorities" (2012, 3). In contrast, on the account suggested here, the notion of preemption is central to our very understanding of trust. For in trusting we see ourselves as having preemptive reasons against taking precautions. What distinguishes those whom we treat as epistemic authorities is

#### 4. Symmetric Speaker-trust

One neglected feature of speaker-trust, distinguishing it from other forms of epistemic reliance, is that the former relations cannot knowingly be symmetric, whereas the latter can be. Consider first epistemic reliance. Two thinkers can knowingly rely upon each other as sources of evidence with respect to a certain proposition. If Reina and Rebecca are both asked to diagnose a patient, they can knowingly rely upon each other in forming their belief: each one, independently, examines the patient and forms a tentative belief about the correct diagnosis before checking with the other. If they then find that they have both arrived at the same diagnosis, they hold on to the belief

---

not the fact that we see their judgment as issuing preemptive reasons for belief, but the strength and scope of these reasons. When a trusted speaker tells me that  $p$ , it might be epistemically responsible for me to treat her saying so as issuing a preemptive reason to believe that  $p$ , but trusting her need not involve thinking that it would be epistemically irresponsible for me to form an opinion on my own consideration of the evidence. In contrast, deference to an epistemic authority involves not merely thinking that I may allow her weighing of the evidence to replace my own, but that it would be irresponsible for me not to treat her judgment in this way (on the difference between speaker-trust and deference to an authority, see also footnote 31).

McMyler does not explicitly employ or address the notion of preemption; nonetheless, his account of epistemic trust is in some ways closer to the one defended here. The main concept underlying McMyler's account is that of a second-personal reason for belief: a reason that, unlike monadic reasons (Thompson 2004), has a fundamental element of address, and that is made available to audiences addressed by a speaker, but not to overhearers. McMyler claims that such second-personal reasons for belief provide the former, but not the latter, with a right of deferral: a right to defer challenges to her testimonially-based beliefs to the speaker on which her belief is based (2011; 2013). While I agree that in trusting speakers we see ourselves as having such a right of deferral, and while this right of deferral fits well within the preemptive-reasons account of trust suggested here, the understanding of this right suggested by my account is quite different than that suggested by McMyler's. First, the account suggested here suggests that in trusting speakers we see ourselves as having not just this right of deferral, but also other rights not to take other kinds of precautions, which are not discussed by McMyler. Thus, this feature of epistemic-trust is an instance of a more general feature of epistemic-trust and of trust more generally. Second, the account suggested here explains why we see ourselves as having this right not by pointing at some second-personal reason for belief, but by arguing that reasons for trust, quite generally, involve second-order reasons against acting for precautionary reasons. One problem with tying this right of deferral (and possibly other rights against taking precautions) with a second-personal reason for belief available only to addressees is that this, *pace* McMyler, makes it difficult to see how a general theory of trust could include such rights in accounting for both testimonial and non-testimonial trust. For it is difficult to see how a general theory of trust can admit that a trustor can have a right of deferral in cases of non-testimonial trust where there is no act of testimonial address, but insist that in the testimonial case, only a trustor who is an addressee can have such a right.

that the diagnosis is correct; otherwise, they suspend judgment. Here Reina and Rebecca knowingly rely on each other as sources of information; in this, each incurs the risk of being led astray (because of incompetence, carelessness, or dishonesty) by her friend. But in an attempt to minimize the risk, each also relies on her own consideration of the evidence. There seems to be nothing particularly bizarre about this policy of symmetric reliance and the fact that it is knowingly adopted by both.

Compare this with Tim and Tom, also both asked to diagnose a patient. Can each of them form his belief about the correct diagnosis on the basis of trust in the other? Can each of them take his friend's word for it? Surely, they cannot do so *knowingly*. Thus, while Tim might believe that the patient has *D* because he believes that Tom believes so, and he trusts Tom, and Tom might believe the same because he believes that Tim believes so, and he trusts Tim, if one of them were to learn the reason for the other's belief, he would no longer be able to maintain his own belief; at least not based on trusting his friend. If Tim learns that Tom's reason for believing that the patient has *D* is that Tim believes so and Tom trusts Tim, Tim might still continue to believe that the patient has *D* based on *other* reasons available to him—he might infer this from the extraordinary combination of symptoms—but in that case, he is not basing his belief upon his trust in Tom.

None of the accounts of trust suggested so far can explain this difference between epistemic reliance and speaker-trust, I would argue, unless the account also involves the claim that reasons for trust are preemptive reasons. According to this claim, trusting a thinker involves seeing his saying or judging that *p* as a reason for not basing one's opinion on *p* on other evidence available to one. On this claim, the problem with symmetric speaker-trust is obvious: if you trust me on *p* and therefore disregard other evidence available to you, and I trust you and therefore disregard other evidence available to me, then both of our beliefs seem to be hanging in midair,



unsupported by any truth-conducive consideration. Once we come to know that this is how our beliefs were formed, this knowledge will surely undermine them. If, however, we are merely relying on each other, we can also each rely on other evidence; in that case, the fact that both of us believe that  $p$ , on the basis of whatever evidence is available to us, provides each of us with further evidence for  $p$ .

How would an account of trust that does not involve the claim that reasons for trust are preemptive reasons explain the difference between speaker-trust and epistemic reliance? Consider Holton's participant-stance account (1994). This account is supposed to explain our alleged ability to decide to trust someone even when we do not believe that she is trustworthy, thus accounting for the possibility of therapeutic trust. But can it explain why Tim and Tom cannot knowingly decide to trust each other about the patient's diagnosis? They can knowingly rely on each other about the diagnosis; they can knowingly adopt the participant's stance towards each other; so why then can they not both knowingly hold a belief about the diagnosis based on trust in each other? Unless it is combined with a preemptive-reasons account, Holton's account seems unable to explain this impossibility. Holton can attempt to explain it by noting that trust, even if it does not require belief, is incompatible with the holding of certain beliefs; that while Tim need not believe that Tom knows that  $p$  in order to trust him, Tim cannot trust Tom if he believes that Tom does not know whether  $p$ .<sup>16</sup> But for this to serve as an explanation, we must still explain why Tim must believe that Tom doesn't know that  $p$  if he knows that Tom's belief is based on symmetrical trust. Unless Tim understands that trust involves disregarding evidence, why couldn't he believe that knowledge was generated through their symmetrical relation of speaker-

---

<sup>16</sup> Holton maintains that while trusting  $B$  to  $\Phi$  does not require believing that  $B$  will  $\Phi$ , it is nonetheless incompatible with believing that  $B$  will not  $\Phi$  (Holton 1994, 71).

trust? The adoption of a preemptive account of trust can allow us to explain the widespread belief that trust does not generate knowledge (Keren 2007), but if trust does not involve such disregard of other evidence, then it is not clear why symmetrical trust should not generate knowledge, just as symmetrical reliance sometimes does.

A similar point applies to other accounts of trust, as involving a belief that the trustee is trustworthy, an affective attitude of optimism, or a non-doxastic disposition to accept the testimony of the speaker. All of these dispositions and attitudes are ones that I can have towards a person at the same time as the person has them towards me. Accordingly, none of these accounts can, by itself, explain the impossibility of known symmetrical speaker-trust. To do so they would need to draw on some belief that must be held by thinkers involved in known relations of symmetrical speaker-trust, and that is somehow incompatible with being in such a relation. Again, the belief that the trusted person does not know seems to be the most plausible candidate. But not every non-doxastic account explains why believing that the speaker does not know is incompatible with believing her. And even those accounts which, like Holton's, are able to explain this appear to be unable to explain why relations of symmetrical trust cannot generate knowledge, unless they accept that trust involves disregarding evidence.<sup>17</sup>

---

<sup>17</sup> An objector might suggest that all that is needed to explain the difference between epistemic reliance and speaker-trust is the contrast between trusting someone for the truth and using someone as a source of information. Any account of trusting-for-the-truth, she may suggest, must say that *A* cannot trust *B* for the truth knowing that *B* is likewise trusting *A* for the truth.

Now I of course agree that any account of trusting-for-the-truth must say this. But why must trusting-for-the-truth satisfy this constraint? Take an account of trusting-for-the-truth that says that *A* trusts *B* for the truth of *p* only if *A* believes that *B* knows that *p* (for a similar account of trust, see Fricker [2006]). Such an account can explain the validity of the constraint only if it can explain why symmetrical trusting-for-the-truth cannot generate knowledge. But why can it not generate knowledge, if trust does not involve disregarding evidence in the way suggested by the preemptive-reasons account? Indeed, Fricker's explanation of why knowledge cannot be generated through (asymmetric)

## 5. Trust, Precaution, Preemption

So far I have made two claims with respect to speaker-trust. The first is that a preemptive-reasons account can explain features of speaker-trust that might initially appear to tell against doxastic accounts of trust—such as the incompatibility of inviting an audience to trust you with criticizing that audience for not weighing evidence. The second is that there are some features of speaker-trust that no other account of trust can explain unless it incorporates the claim that reasons for trust are preemptive reasons. I now want to argue that similar claims are true of trust in general.

No account of trust can be complete, I claim, unless it incorporates the idea that reasons for trust are preemptive reasons. Accounts of trust in terms of a mental attitude—belief, affective attitude, participant's stance—that do not also adopt the idea that trust involves responding to preemptive reasons imply that we can trust someone to  $\Phi$  while at the same time taking precautions against her failing to  $\Phi$ . Even if we often do not take such precautions, they are not incompatible with such mental attitudes. They are, however, incompatible with trust.<sup>18</sup>

---

trust does not work for symmetrical trust. In the asymmetric case, the very idea of knowledge generation requires trust to be based on a false belief that the speaker knows; according to Fricker, this explains why such trust does not generate knowledge. But in symmetrical trust, the idea that trust generates knowledge does not entail that one knows on the basis of a false belief. Here, then, the claim that symmetrical trust does not generate knowledge remains unexplained. After all, it has been suggested that testimony is sometimes a generative source of knowledge (Lackey 1999, 2006; Graham 2000). If, as some have suggested, trust, unlike other ways of relying on testimony, does not generate knowledge (Fricker 2006; Keren 2007), this is something that an account of trust must explain if it is to explain the difference between symmetrical trust and symmetrical reliance.

Finally, our objector may try to explain why A cannot trust B for the truth knowing that B is likewise trusting A by noting that trust-for-the-truth involves taking the trustee as an authority. But this explanation, even if successful, does not help our objector. For it is of the essence of authority, practical or epistemic, that it issues preemptive reasons (Zagzebski 2012, Raz 1990a), so this is an explanation that supporters of the preemptive reasons account will be happy to endorse.

<sup>18</sup> I do not claim that other accounts cannot *accommodate* the idea that one does not trust  $B$  to  $\Phi$  if one takes precautions against  $B$  not  $\Phi$ 'ing, but that they do not *entail* it. Therefore, unless they incorporate the claim that reasons for trust are preemptive reasons, they are incomplete.

A person who believes that  $p$  can nonetheless take every precaution against the possibility that her belief is false. She can do so because the stakes are so high that, despite her confident belief, taking precautions against this seemingly improbable possibility appears to her to be rational; or she can do so when she herself finds doing so irrational, as in the case when one knows something to be safe but is nonetheless swept by recalcitrant fear (Benbaji 2013). Similarly, we can have an affective attitude of optimism about a trustee's goodwill but take every precaution against being let down—perhaps because we believe that this optimism is misplaced. The same is true when we rely on a person while adopting the participant's stance towards her.

Trust, however, is not compatible with such excessive precautions nor with seeing oneself as having no reason against taking precautions. Thus, a shop owner might leave her employee alone in the shop with a significant amount of money in the till while she goes out on some important errand. But if, before she leaves, she turns on the CCTV camera to monitor the employee's movements, then she does not really trust him. Even if she believes that he is trustworthy, or is optimistic that he will not steal, and turns on the cameras just as a precaution, it would not be correct to say of her that she trusts him not to steal. The extent of our trust is at least partially determined by the degree to which we see ourselves as having a preemptive reason against taking precautions and the degree to which we act accordingly.<sup>19</sup>

One might admit that trust requires not acting in certain ways (not operating CCTV cameras, not checking on the kids during intermission...) but still question whether trust indeed involves responding to preemptive reasons. Preemptive reasons

---

<sup>19</sup> As I note in section 7, trust comes in degrees for at least two reasons: First, like other beliefs, the belief that a person is trustworthy which is necessary for trust, comes in degrees. Second, a person can be believed to be more or less trustworthy. Both of these are reflected in the kind of precautions we see ourselves as having reason not to take, and in the strength of these reasons. See also footnote 31.

for action are not merely reasons against acting in certain ways; they are reasons against acting for certain reasons (Raz 1990a). But as long as the employer does not operate CCTV cameras, does it matter what his reasons are for not doing so? Does it matter that the employer is responding to *preemptive* reasons?

I want to insist that unless we are responding to preemptive reasons, we are not trusting. Sometimes we see ourselves as having every reason to take precautions but fail to take them nonetheless: not because we see a reason not to but because something makes us neglect doing what we take ourselves to have every reason to do. This is not trusting. If the employer is so nervous about the new employee that she just forgets to turn on the CCTV cameras, her failing to do so has nothing to do with her trusting him. To exhibit trust, one must see oneself as having reason *not* to take precautions.

Moreover, trust requires not just any reason against acting; it requires preemptive reasons. If I call the babysitter not to check that everything is OK but just to tell her how grateful we are for her help, my calling exhibits no lack of trust. If the employer turns on the cameras not as a precaution against possible theft, but simply because she wants to use the recording in training future employees, her employee cannot complain that she does not trust him. These actions do not exhibit lack of trust, because while performing an action that often is preformed as a precaution, the agent is not actually taking a precaution, because she is not acting for the kind of reason that would render the action precautionary. Acting in certain ways amounts to taking a precaution only when the action is performed for a certain reason, namely to guard against a certain risk. Accordingly, reasons against taking precautions are not merely reasons against acting in certain ways; they are reasons against acting *for certain*

*kinds of reasons*; they exclude, or preempt, certain reasons, namely, precautionary reasons. Hence they are preemptive reasons.<sup>20</sup>

## 6. The Doxastic Account Reconsidered

I have claimed that no account of trust that does not incorporate a preemptive-reasons account can explain some of the distinctive features of trust. I now want to argue that this may apply to some of the very features on which objections to doxastic accounts of trust have been based. A preemptive-reasons account suggests a better explanation of some of these features than that provided by non-doxastic accounts, and it does so without committing us to a non-doxastic account of trust.

As we have seen, one type of objection to doxastic accounts of trust emerges from the resistance of trust to evidence. Another comes from the fact that trust is undermined by excessive reflection. But how are we to explain the fact that trust has both of these features? If trust involves responding to preemptive reasons against taking precautions, then we should expect trust to be both resistant to counter-evidence and to be undermined by extensive reflection on relevant evidence. Both features appear to be manifestations of a single underlying feature of trust: that in trusting someone to  $\Phi$  I must respond to preemptive reasons against my taking precautions against her not  $\Phi$ 'ing. We have already noted that when we rely on someone for a belief, considering relevant evidence is a main precaution that we can

---

<sup>20</sup> The main point I want to insist on here is that reasons for trust are preemptive reasons in the sense that they exclude or preempt other reasons. That is, they are second-order reasons against acting (or believing) on certain first-order reasons. However, as noted above, the term "preemptive reason" is used by some as equivalent with the Razian notion of "protected reasons," and on the account suggested here, reasons for trust can also be described as protected reasons. For if we trust someone to  $\Phi$ , we not only rely on her to  $\Phi$ , we rely while being prepared not to take precautions against the possibility that she will not  $\Phi$ . That is, we rely while being prepared not to act on possible reasons for not relying. Thus reasons for trust are a combination of first order reasons to rely in certain ways, and of second-order (exclusionary) reasons against acting for certain conflicting reasons. Thus, they are protected reasons.

take in order to guard against the risk of being let down. A similar point applies when we rely on others in other ways. After all, other precautions we can take—installing CCTV cameras, calling the babysitter to check on things—serve as precautions precisely because they can provide us with evidence about the risk of being let down. To guard against that risk, we can reflect on the evidence that has suggested to us that we will not be let down or we can remain attuned to relevant counter-evidence. Reasons against taking precautions, then, are not only reasons against acting in ways that might produce such evidence but also reasons against reflecting on or being attuned to such evidence.<sup>21</sup> Thus, if we insist on reexamining our grounds for trust, we are not fully responding to such preemptive reasons. And if we do respond to such reasons, we will shift our attention away from such evidence, developing a certain resistance to it.<sup>22</sup>

---

<sup>21</sup> More accurately, reasons against taking precautions are reasons against reflecting on such evidence in as much as such reflection is motivated by precautionary reasons or is apt to motivate a precautionary attitude. For neither reflecting on our grounds for trust, nor operating the CCTV cameras amounts to taking precautions, if they have nothing to do with guarding against relevant risks. Thus, for instance, reasons against taking precautions are reasons against reflecting on our past interactions with the trustee, when such reflection is motivated by our desire to guard against the risk of being let down, but not when we know that such reflection will not affect our action, and merely reflect on them while being swept by a wave of nostalgia. In this respect, the kind of preemptive reasons which are central to our understanding of trust are similar to the kind of preemptive reasons central to Raz's account of authority and of authoritative rules. For the latter too, while described as reasons for disregarding certain reasons, do not exclude engaging in thought about the latter reasons, so long as we know that our reflection will not affect our action (Raz, 1990b, 183-4). I am grateful to an anonymous reviewer of this journal for pressing me on this point.

<sup>22</sup> It might be objected, that the preemptive-reason account explains why trust is incompatible with reflection, but not why such reflection undermines trust. In cases where we have good evidence for the trustee's trustworthiness, why would such reflection not result in the restoration of trust, even if it (temporarily) displays lack of trust? The reason for thinking that even in such cases trust would often not be restored has to do with the fact that trustworthiness is partially a matter of responsiveness to the trustor's trust and dependence on the trustee. Excessive reflection, because it displays lack of trust, may negatively affect the trustee's willingness to react to the trustor's reliance on her, even if trust is restored. Moreover, this is something the trustor can anticipate. Accordingly, even if initially trust was based on strong evidence supporting the trustee's trustworthiness, reflection on this evidence might result not in the restoration of trust, but in the restoration of the belief that the trustee would have been properly responsive to the trustor's trust, if the trustor had not displayed lack of trust, conjoined with the belief that now that lack of trust has been displayed, the trustee would not similarly respond to

Of the non-doxastic accounts currently on the market, an affective-attitude account appears to be the one best suited to explaining the two features mentioned above.<sup>23</sup> Trust's resistance to evidence has been cited as a central reason for favoring such an account of trust (Jones 1996; McLeod 2002): emotions, like trust, are said to be resistant to evidence and give rise to beliefs that are resistant to evidence. Yet despite this similarity, an affective-attitude account of trust does not provide a satisfying explanation for the kind of resistance exhibited by trust. When we trust a person, we do not merely tend not to pay attention to certain pieces of evidence that might seem to suggest that she may let us down. We also see ourselves as having a *reason* against paying attention to such evidence. A parent, who trusts a babysitter but happens upon evidence that may raise questions about the quality of the care, might dismiss this evidence, citing the babysitter's trustworthiness as her reason for doing so: "Yes, I heard little Johnny crying. But you know we can trust Sarah, so let's just watch the show and stop thinking about that." In contrast, while emotions cause us to focus on a partial field of evidence, once the partiality of our focus has been brought to our attention, our emotions do not seem to justify dismissing counter-evidence to which we had not been attuned. My fear of little Snoopy might make me disregard our numerous harmless encounters, but once I am reminded of them, I will not see my fear as grounds for dismissing this evidence of Snoopy's harmlessness.

---

restored trust. I am grateful to an anonymous reviewer of this journal for pressing me on this point.

<sup>23</sup> An account of trust as involving the adoption of the participant's stance is supposed to explain some features of trust, such as the phenomenon of therapeutic trust (Holton 1994), but does not seem to explain these two, while an account of trust as a non-doxastic disposition to rely in certain ways does not appear to be in the business of explaining features of trust. Describing trust as a disposition to act or form beliefs in certain ways does not explain why trust has the features that it has.



When we are made aware of tensions between our trust and available evidence, this can give us a reason to dismiss the evidence or to avoid considering it. This seems to support the claim that trust's resistance to evidence and the fact that trust is undermined by reflection on relevant evidence are two manifestations of the same underlying fact, as suggested by the preemptive-reasons account. An affective-attitude account, meanwhile, because it does not portray the resistance of trust to evidence as reason-based, has trouble explaining why trust is undermined by reflection on the evidence and fails to explain the relation between the two noted features of trust. Emotions, in addition to their resistance to evidence, are also resistant to considerations of the evidence in the sense that they typically withstand such consideration.<sup>24</sup> Trust, in contrast, while resistant to evidence, is not at all resistant to consideration of the evidence, which tends to undermine trust. An affective-attitude account seems to face a serious challenge in its attempt to explain this combination of features.

Thus, two features of trust that have seemed to support non-doxastic accounts do not after all tell against a doxastic account of trust, unless a doxastic account is incompatible with the idea that reasons for trust are preemptive reasons. But this is hardly the case, as I will suggest below.

## **7. Trusting without Believing?**

The suggestion that trust is a distinctive form of reliance partially because it involves responding to reasons that preempt taking precautions, can in principle be combined with various other accounts of trust, doxastic or non-doxastic. Indeed, there are reasons to think that it must be combined with, or derived from, some other account of

---

<sup>24</sup> Even recalcitrant emotions, which appear to a subject to be in tension with evidence available to her, characteristically survive reflection (Benbaji 2013).

trust. For even if it is correct that no account of trust is complete without the idea that reasons for trust are preemptive reasons, it is doubtful whether this idea provides us with an adequate account of trust: arguably, that I avoid taking precautions because I see myself as having preemptive reasons against taking precautions is a necessary but not sufficient condition for trust.

Does the idea that reasons for trust are preemptive reasons fit best with a doxastic or a non-doxastic account of trust? My claim that some objections to doxastic accounts are neutralized once we accept this idea does not settle the question. Nor does the idea neutralize all objections to doxastic accounts of trust. In particular, it does not help much with the objection from therapeutic trust. Therefore, to complete my defense of a doxastic account of trust, I would like to address this objection by considering some relevant cases. These, I suggest, support a doxastic account of trust as best suited to go along with a preemptive-reasons account of reasons for trust.

First, let us consider why responding to preemptive reasons against taking precautions is not a sufficient condition for trust. Consider the case of Captain Jones, a military officer, worried about relations of distrust within a unit under his command, who therefore orders his sergeant to act as if he trusts Private Smith. If the sergeant recognizes Captain Jones's authority, he will see the order as generating preemptive reasons for him not to take precautions against the risk of being let down by Private Smith. Nonetheless, even while responding to these reasons, the sergeant might have no trust at all in Private Smith. As long as he sees himself as responding to preemptive reasons the source of which is an authoritative command rather than anything he ascribes to Private Smith, he hardly seems to trust the private.

Thus, the fact that we see ourselves as having preemptive reasons against taking precautions is not sufficient for trust; for it to be trust, we must see our reasons against taking precautions as having *the right kind of source*. But this suggests that

what the sergeant lacks is not an affective attitude of optimism about Private Smith, nor a disposition to rely on him, nor a readiness to accept the participant's stance towards him. For if what I have said in the previous section is correct, then none of these can be the source of the kind of preemptive reasons we see ourselves as having when we trust someone, namely reasons to dismiss counter-evidence or not to consider the evidence.<sup>25</sup> In contrast, we *can* have such reasons in virtue of our having a belief with appropriate content about the trustee: if I have reason to believe that she is honest, competent, and in a better position to judge than I am, this may provide me with reasons for not basing my belief on my own weighing of the evidence (Zagzebski 2012); if I have reason to believe that she has goodwill and is the kind of person who will respond to my dependence on her by being particularly careful not to let me down, this can be reason enough not to take such precautions: not to monitor her behavior, nor to be attuned to evidence otherwise produced and which might be relevant to the possibility of her not doing what I depend on her to do. Thus, a doxastic account seems to be the most plausible, given the kind of preemptive reason which we see ourselves as having when we trust someone and given that trusting her requires that we see these reasons as having the right kind of source.

Our discussion therefore leads us to adopt the following (partial) analysis of trust: *A* trusts *B* to  $\Phi$  only if *A* believes that *B* is trustworthy,<sup>26</sup> such that in virtue of *A*'s belief about *B*'s trustworthiness, *A* sees herself as having reason to rely on *B*'s  $\Phi$ 'ing without taking precautions against the possibility that *B* will not  $\Phi$ , and only if

---

<sup>25</sup> This is not to claim that non-doxastic attitudes cannot provide us with preemptive reasons, but rather that the attitudes that are central to non-doxastic accounts of trust do not provide the *kind* of preemptive reasons, such as reasons against being attuned to contrary evidence, that are characteristic of trust.

<sup>26</sup> Where *B*'s trustworthiness with respect to  $\Phi$  is to be understood in terms of relevant competence, goodwill, and responsiveness to *A*'s dependence on her.

*A* indeed acts on, or is responsive to, reasons against taking precautions.<sup>27</sup> In particular, in the case of speaker-trust, *A* speaker-trusts *B* on *p* only if *A* believes that *B* is trustworthy with respect to *p*, such that in virtue of this belief, *A* sees herself as having a reason to believe *B*'s testimony without taking precautions against the possibility that *B*'s testimony is false, and only if *A* is responsive to this reason.<sup>28</sup>

But then what are we to make of the objection to doxastic accounts from therapeutic trust? The idea that reasons for trust are preemptive reasons does not help against this objection as it does against other objections studied earlier, for it may appear that the possibility of trusting someone without believing that she is trustworthy is implicit in the very idea of therapeutic trust: if we can trust someone in order to promote her trustworthiness, doesn't it follow that we can trust someone even if we do not believe her to be trustworthy?

This, however, is too quick: trust and trustworthiness come in degrees, as do beliefs. Parents may trust their teenage daughter (to a degree), believing (to a degree) that she is trustworthy (to a degree), while also believing that thanks to their trust, she may become more trustworthy. The question is whether the more doubts we have about a person's trustworthiness, the less the degree of our trust in her. Proponents of

---

<sup>27</sup> Sometimes a trustor may see herself as having reason for not taking precautions, but not act on this reason because she has no opportunity to take relevant precautions. Nonetheless, she may be responsive to this reason if she is willing not to take certain precautions, should she have the opportunity to take them

<sup>28</sup> While it is the belief that the trustee is trustworthy that is central to trust on the account suggested here, full-fledged trust also involves the belief that the trustee will do what she is trusted to do. Thus, as suggested in section 2, if *B* tells *A* that *p*, and *A* fully trusts *B* to tell her the truth, then *A*'s strong belief that *B* is trustworthy (combined with *A*'s disregard of other evidence that she may have), would lead *A* to believe that *B*'s testimony is true. Nonetheless, when one does not fully trust the trustee, but only trusts her to a degree, one may believe, to a degree, that the trustee is trustworthy, without believing to the same degree that she will do what she is trusted to do. This is a result of the fact that being trustworthy with respect to  $\Phi$ 'ing is compatible with failing to  $\Phi$ , and the fact that when trust is not complete, one may be responsive, to some extent, to evidence suggesting that the trustee may not do what she is trusted to do.

a doxastic account of trust would have to insist on a positive answer to this question (Hieronymi 2008).<sup>29</sup> Do examples of therapeutic trust suggest otherwise?

The standard example—parents who trust their teenage daughter to look after the house in order to promote her trustworthiness—provides no clear-cut answer. First, in describing the parents as attempting to promote her trustworthiness by trusting her, it seems that we might be attributing beliefs to them—about her goodwill, moral competence, and responsiveness to their trust—that amount to ascribing a degree of trustworthiness to her.<sup>30</sup> Secondly, to the extent that the parents have grave doubts about their daughter's trustworthiness, it might be asked whether they *really do* trust her or merely act *as if* they do.

Because standard examples may not allow us to test our intuitions about this, we should look at cases that depart from such standard examples in two significant ways: first, cases in which one's reasons for trusting a person believed to be lacking in trustworthiness emerge not from one's beliefs about her but rather from reasons provided by a third party (as with Captain Jones); and second, cases in which the trustee is in fact very trustworthy and can therefore, unlike the teenager in the standard example, complain if she is not *really* trusted. It seems that our intuitions regarding both kinds of cases conform to what a doxastic account would lead us to expect.

Consider Captain Jones again, worried this time about the absence of relations of trust within his own family. His wife, Anne, does not trust his mother: Anne has

---

<sup>29</sup> More accurately, proponents of doxastic accounts are committed to two claims: first, that one does not trust a person at all if one in no way believes her to be trustworthy; and second, that other things being equal, the greater one's doubt regarding her trustworthiness, the lesser one's trust (Hieronymi 2008).

<sup>30</sup> Indeed, prominent supporters of non-doxastic accounts have provided accounts of trustworthiness according to which a trustworthy person is someone who is motivationally responsive to the fact that she is counted upon (Faulkner 2011, 148; Jones 2012).

grave doubts about Mother Jones's goodwill, about her moral competence, etc. Moreover, she has grave doubts whether Mother Jones will ever be able to change. She knows that her husband is hurt by her lack of trust in his mother and would very much like her to trust Mother Jones: not just to act *as if* she trusted Mother Jones, but to *really* trust her. Indeed, because she would like to make him happy, she too would like to trust Mother Jones. But while he believes his mother to be trustworthy, she does not. Is there a way for her to make herself trust Mother Jones that does not involve at least partially removing her doubts about her trustworthiness? She can surely act *as if* she trusted Mother Jones, while harboring these doubts. But can she really come to trust her without forming the belief that Mother Jones is trustworthy, at least to a degree? It is not at all clear that she can.

Consider next the case of the Barneses. Mr. Barnes is a faithful husband, a model of loyalty and trustworthiness, but Mrs. Barnes is of the suspicious kind and has always doubted his faithfulness. Her doubts almost led to the collapse of their marriage, before she finally agreed to seek counseling. Now she no longer makes scenes; she acts just like a trusting wife would. Moreover, as a result of the treatment she underwent, she has developed an affective attitude of optimism regarding Mr. Barnes's goodwill and moral competence. Nonetheless, she still has grave, unexpressed doubts about his goodwill and his ability to remain faithful to her if faced with an opportunity to betray her. Deep down, she believes that her optimism about his moral competence is unfounded and that her continuing doubts about him are justified. If Mr. Barnes were to discover that she has such doubts about him, could he not complain to her that she does not *really* trust him? And if he did, could Mrs. Barnes respond by saying that she in fact really does trust him, but only has grave doubts about his trustworthiness? Surely she could not.

Thus, the argument from therapeutic trust is weak. It appeals to an alleged intuitive judgment that we can trust without believing, but on closer inspection—once we are careful to distinguish between a limited belief and no belief in a person's trustworthiness, and between cases in which one really trusts and cases in which one merely acts *as if* one trusted—it becomes very unclear that we actually have such an intuition.

### **8. The Rationality of Trust and of Trust-Based Beliefs**

As Jones (1996) has noted, the conditions that a relation of trust must satisfy in order to be rational depend on the mental attitude involved in trust. Moreover, we should add, the relation between the rationality of trust and the justification of trust-based beliefs may also depend on this. What, then, are the implications of my defense of a doxastic account of trust for the conditions for the rationality of trust and of trust-based beliefs?

Accepting the account defended here will not in itself tell us what conditions must be met for trust to be rational or for trust-based beliefs to be (epistemically) justified. Consider the question of epistemic justification: Different theories of justification would suggest different conditions that must be met for a trust-based belief to be justified. To answer the question, therefore, we would have to adjudicate among these different theories, which we cannot do here. Nonetheless, we can describe the rough outlines of a core idea that different epistemological theories could all employ to explain how and when trust-based beliefs can be justified.

Trusting a speaker, and thus responding to preemptive reasons for belief, makes one's belief less sensitive to evidence available to one. However, it also makes one's belief more sensitive to evidence available to the speaker. As a result, when speakers have an epistemic advantage over audiences, audiences' epistemic position

will often be no worse, and sometimes better, if they trust speakers instead of weighing their testimony against all evidence available to them. The clearest case is the one in which the speaker is not only honest and competent but also a known expert on the issue. But even if the speaker does not have greater expertise, but only enjoys "positional advantage" over audiences (Williams 2002), the epistemic position of audiences may still be better, or at least no worse, if they trust the speaker instead of weighing her testimony merely as one piece of evidence.<sup>31</sup>

Different epistemological theories will employ this core idea in different ways and will therefore suggest different conditions for the justification of trust-based beliefs. A reliabilist can understand the idea of an epistemic advantage enjoyed by a speaker, and our talk of one's belief being more sensitive to the speaker's evidence, in terms of the reliability of available belief-forming mechanisms.<sup>32</sup> An internalist can make sense of the core idea by saying that when audiences have reasons to believe that the speaker is honest and competent, and has access to a better body of evidence, they also have reasons to believe that by trusting her, their belief will be more likely to fit the speaker's better evidence. Alternatively, epistemologists committed to the

---

<sup>31</sup> There are important differences between the two kinds of cases. An expert speaker typically has advantage over her lay audiences both in terms of the evidence available to her, and in terms of her ability to evaluate this evidence; a non-expert speaker with positional advantage over her audience normally enjoys only the former kind of epistemic advantage. As a result, the preemptive reasons issued by the testimony of these two kinds of speakers differ significantly in terms of both scope and strength. The judgment of a known expert preempts consideration of more evidence than that of a non-expert with positional advantage. Moreover, the preemptive reasons issued by an expert's testimony often make it epistemically irresponsible for us to base our belief on our own weighing of the evidence, whereas in the case of a non-expert speaker, such reasons often permit, but do not require, not weighing certain evidence.

<sup>32</sup> Zagzebski (2012), following Raz (1986), suggests an argument for treating an expert's judgment as a preemptive reason for belief (rather than as a reason for belief to be added to all others), which, if successful, should be convincing to the reliabilist. For the argument is supposed to establish the claim that under certain conditions, treating the expert's judgment as a preemptive reason would improve our "track-record" (Zagzebski 2012, 114) or reduce our "rate of mistake" (Raz 1986, 68), when compared with other ways of treating her judgment.



"inheritance model" (McMyler 2011) can make sense of the idea by saying that by trusting a speaker, audiences allow their beliefs to inherit the evidential support enjoyed by the speaker's belief.

While different epistemological theories have different implications for the conditions under which speaker-trust, on the account defended here, would lead to the formation of justified beliefs, I want to propose two claims that should be accepted under any plausible theory: first, that the rationality of the trust put in a speaker is neither necessary nor sufficient for the epistemic justification of beliefs based on that trust; and second, that the justification of trust-based beliefs must ultimately depend on the justification of beliefs that are not trust-based.

The first claim follows from the fact that while both speaker-trust and trust in general require belief in the trustworthiness of the trustee, this belief being justified is not a necessary condition for the rationality of trust but *is* a necessary condition for the justification of trust-based beliefs. Some philosophers have assumed that accepting an account of trust that involves a belief in the speaker's trustworthiness as a central component commits us to saying that trust is rational only if the belief required for trust is itself justified (Jones 1996). This is a mistake. As already noted, trust, though it requires a belief, involves more than merely holding a belief.<sup>33</sup> Hence, the standards of appraisal applicable to trust need not be exclusively epistemic. Moreover, there are reasons to think that trusting may be rational even if the belief required for trust is not epistemically justified. Compare: It might sometimes be rational for a person to act in ways that will make her form a belief even if the resultant belief itself would not be epistemically justified. If a cancer patient can raise his chances of survival by making himself believe, contrary to the evidence, that he will survive his

---

<sup>33</sup> See footnote 4.

illness, then forming that belief might be rational, all things considered, even if the belief itself is unjustified. Similarly, trust might be rational, all things considered, even if the belief in the trustworthiness of the trusted person is epistemically unjustified. Sometimes the likely consequences of trust can support trusting a person even if the evidence does not suggest that she is trustworthy. If a young man trusts his first love to be faithful, his belief in her trustworthiness might be epistemically unjustified, yet if holding that belief is necessary in order for love to flourish, that trust might nonetheless be rational.<sup>34</sup>

Matters are different with epistemically justified trust-based beliefs. Trust-based beliefs cannot be epistemically justified unless belief in the trustworthiness of the speaker is justified, for trust-based beliefs are based upon belief in the speaker's trustworthiness. As almost all accounts of epistemic justifications, internalist or externalist, agree, when one belief is based on another, or is the product of a belief-dependent process involving another belief (Goldman 1979), the latter belief must be justified if the former is to be.

These two claims, taken together, lead to the conclusion that rational trust in a speaker is neither necessary nor sufficient for the epistemic justification of the resulting trust-based belief. If justified belief in the trustworthiness of a speaker is a necessary condition for the epistemic justification of trust-based beliefs, but not for the rationality of trust, then the rationality of trust is not a sufficient condition for the

---

<sup>34</sup> It might be objected that, where a belief in the trustworthiness of a person is unjustified, it would be better not to hold the belief, and yet to enjoy the possible fruits of trust by acting as if one trusted. However, this is not always possible. Trust involves doing things—such as focusing our attention in certain ways—that we often cannot do unless we believe in the trustworthiness of a person. Moreover, trusting relationships are often valuable in themselves, and we cannot engage in them unless we really trust our partner to the relationship. Even if Mrs. Barnes's doubts about her husband's faithfulness are not manifested in any way in her behavior, they still affect the couple's relationship. As long as she harbors them, she is not engaged in a trusting relationship, even if she acts as if she were. Therefore, our ability to enjoy the fruits of trust without really trusting is limited.

epistemic justification of trust-based belief. Moreover, the kind of considerations that have led us to this conclusion suggest that justified belief in the trustworthiness of the trustee is not sufficient for the rationality of trust and therefore that rational trust is not necessary for epistemically justified trust-based belief.

We must also conclude that ultimately, on pain of regress, the justification of trust-based beliefs depends on the justification of non-trust-based beliefs. Belief in the trustworthiness of a speaker must not be trust-based or, if trust-based, must ultimately be based on a belief, not trust-based, in the trustworthiness of another speaker. One might object that this conclusion follows only for a foundationalist committed to a linear conception of epistemic justification. But even a coherentist conception of justification cannot avoid this conclusion, because a belief that is justified because it coheres well with the believer's own body of belief is arguably not a trust-based belief.

We may seem to have come to the conclusion that there is nothing epistemically distinctive about trust-based beliefs because their justification is reducible to that of non-trust-based beliefs. But I have not claimed that the one is reducible to the other; instead, I have argued for the weaker claim that trust-based beliefs depend for their justification on non-trust-based beliefs. More importantly, my argument for the latter claim emerges from the defense of a doxastic account of trust resting on the claim that reasons for trust are reasons of a distinctive kind. This discussion, rather than reducing the question of the justification of trust-based beliefs to familiar questions about more widely discussed sources of epistemic justification

should draw our attention to a distinctive kind of epistemic reason: preemptive reasons for belief.<sup>35</sup>

### **Acknowledgements**

Research on this paper was generously funded by the Israel Science Foundation (grant no. 714/12). An earlier version of this paper was presented at the International Conference on Theoretical and Rational Rationality at the Israel Institute of Advanced Studies. For helpful comments, I am grateful to participants at the conference, to the editors and anonymous reviewers of this journal, and to Iddo Landau, Ariel Meirav, Collin O'neil, and Saul Smilansky. I am especially grateful to Joseph Raz.

---

<sup>35</sup> See Zagzebski (2012) for an important discussion of such reasons and of their neglect within much of modern philosophy.

## References

- Adler, J. E. (1994). Testimony, trust, knowing. *Journal of Philosophy* 91 (5), 264–275.
- Baier, A. (1986). Trust and antitrust. *Ethics* 96 (2), 231–260.
- Baker, J. (1987). Trust and rationality. *Pacific Philosophical Quarterly* 68, 1–13.
- Benbaji, H. (2013). How is recalcitrant emotion possible? *Australasian Journal of Philosophy* 91 (3), 577–599.
- Darwall, S. (2010). Authority and reasons: exclusionary and second-personal. *Ethics* 120 (2), 257–278.
- Elster, J. (2007). *Explaining social behaviour: more nuts and bolts for the social sciences*. Cambridge: Cambridge University Press.
- Faulkner, P. (2007). On telling and trusting. *Mind* 116 (464), 875–902.
- Faulkner, P. (2011). *Knowledge on trust*. Oxford: Oxford University Press.
- Faulkner, P. (2012). Replies. *Abstracta* 6, 117–137.
- Fricker, E. (2006). Second-hand knowledge. *Philosophy and Public Affairs* 73 (3), 592–618.
- Frost-Arnold, K. (forthcoming). The cognitive attitude of rational trust. *Synthese*.
- Goldman, A. I. (1979). What is justified belief? In G. Pappas (ed.), *Justification and knowledge*. Dordrecht: Reidel.
- Graham, P. J. (2000). Conveying information. *Synthese* 123 (3), 365–392.
- Hieronimi, P. (2008). The reasons of trust. *Australasian Journal of Philosophy* 86 (2), 213–236.
- Holton, R. (1994). Deciding to trust, coming to believe. *Australasian Journal of Philosophy* 72(1), 63–76.
- Jones, K. (1996). Trust as an affective attitude. *Ethics* 107 (1), 4–25.
- Jones, K. (2004). Trust and terror. In P. DesAutels and M. U. Walker (eds.), *Moral psychology: feminist ethics and social theory*. Lanham, MD: Rowman & Littlefield.
- Jones, K. (2012). Trustworthiness. *Ethics* 123 (1), 61–85.
- Kappel, K. (forthcoming). Believing on trust. *Synthese*.
- Keren, A. (2007). Epistemic authority, testimony and the transmission of knowledge. *Episteme* 4 (3), 368–381.

- Lackey, J. (1999). Testimonial knowledge and transmission. *Philosophical Quarterly* 50 (197), 471–490.
- Lackey, J. (2008). *Learning from words: testimony as a source of knowledge*. Oxford: Oxford University Press.
- McLeod, C. (2002). *Self-trust and reproductive autonomy*. Cambridge, MA: MIT Press.
- McLeod, C. (2011). Trust. *The Stanford encyclopedia of philosophy* (Spring 2011 edition), ed. Edward N. Zalta.  
<http://plato.stanford.edu/archives/spr2011/entries/trust>. Accessed 3 September, 2012.
- McMyler, B. (2011). *Testimony, trust, and authority*. Oxford: Oxford University Press.
- Moran, R. (2005). Getting told and being believed. *Philosophers' Imprints* 5, 1–29.
- Pettit, P. (1995). The cunning of trust. *Philosophy & Public Affairs* 24 (3), 202–225.
- Raz, J. (1986). *The morality of freedom*. Oxford: Clarendon Press.
- Raz, J. (1990a). Authority and justification. In J. Raz (ed.), *Authority*. Oxford: Basil Blackwell.
- Raz, J. (1990b). *Practical reasons and norms*, 2nd ed. Princeton: Princeton University Press.
- Thompson, M. (2004). What is it to wrong someone? A puzzle about justice. In R. Jay Wallace, Philip Pettit, Samuel Scheffler and Michael Smith (eds.), *Reason and value: themes from the moral philosophy of Joseph Raz*. Oxford: Oxford University Press.
- Williams, B. (2002). *Truth and truthfulness: an essay in genealogy*. Princeton: Princeton University Press.
- Zagzebski, L. (2012). *Epistemic authority: a theory of trust, authority and autonomy in belief*. Oxford: Oxford University Press.