The Cognitive Attitude of Rational Trust

Karen Frost-Arnold

Hobart & William Smith Colleges

Abstract

I provide an account of the cognitive attitude of trust that explains the role trust plays in the planning of rational agents. Many authors have dismissed choosing to trust as either impossible or irrational; however, this fails to account for the role of trust in planning. *A* can have therapeutic, coping, or corrective reasons to trust *B* to ϕ, even in the absence of evidence that *B* will φ. To accommodate such types of trust, without accepting doxastic voluntarism, requires an account of the cognitive attitude of trust broader than belief alone. I argue that trust involves taking the proposition that someone will do something as a premise in one's practical reasoning, which can be a matter of believing or accepting the proposition. I defend this account against objections that it (i) provides insufficient rational constraints on trust, (ii) conflates trust and pretense of trust, and (iii) cannot account for the rationality of back-up planning.

1 Introduction

Interpersonal trust is a rich concept for ethicists and epistemologists, because it focuses attention on our interdependency and vulnerability to each other. This vulnerability makes

trustworthiness both a moral and epistemic virtue. But trust has proven challenging to analyze, because one can trust for a variety of reasons. Consider trusting someone to tell the truth. One might trust another's testimony because one has good evidence to believe inof her sincerity, because one hopes that one's trust will inspire greater honesty in the speaker, because trusting will reduce one's anxiety or cognitive load, or because one wishes to correct for one's own prejudice against the speaker. An account of trust broad enough to encompass all these types of trust has eluded philosophers, in part, because several difficult questions must be answered. What is the cognitive attitude involved in trust? How is trust integrated with practical reasoning and action? What, if any, are the rational constraints on trust?

These questions are related. Questions about the rationality of trust have proven particularly vexing because trust is not always responsive to evidence. Sometimes the evidence that someone will do somethingof trustworthiness compels us to trust her to do it, but, as I will argue, other times we choose to trust someone to do something without good reason to believe that they will do so. This raises doubts about the cognitive attitude involved in trust. Belief is commonly thought to be an involuntary attitude that aims at truth; however, trust may be chosen for non-epistemic reasons. Thus, if the common view of belief is correct, the cognitive attitude of trust is not always belief.

While some authors have suggested that some types of trust involve a cognitive attitude other than belief (Holton, 1994; Faulkner 2007, 2011), none have provided a sustained analysis of the cognitive attitude of trust that (a) explains what trusting belief and trust without belief have in common, and (b) successfully responds to pressing objections to such a view. This paper provides such an account: trust involves taking the proposition that the trusted will act as expected as a premise in one's practical reasoning. Sometimes this involves the cognitive attitude of belief, while other times it involves the attitude of acceptance. This account grounds

the cognitive attitude of rational trust in its role in planning.

I proceed as follows. First, I motivate the need for this account by showing that trust can be chosen for non-epistemic reasons. I argue that while previous authors have attended to one type of chosen trust, there are actually at least three types of chosen trust that present problems for the claim that trust requires belief. Second, I present my account of the cognitive attitude of trust and use it to explain the three types of chosen, non-evidential trust. Third, I defend the account from objections that it (i) provides insufficient rational constraints on trust, (ii) conflates trust and pretense of trust, and (iii) cannot account for the rationality of back-up planning.

2 Terminological Distinctions

Two terminological distinctions are necessary. First, interpersonal trust, in the sense discussed here, is a three-part relation in which A trusts B to ϕ. This type of trust always involves a truster, A, a trusted, B, and an action, ϕ, that the truster trusts the trusted to perform. Thus, I will not be concerned with questions about how generally trusting people are (this would be to analyze trust as a monadic property), nor will I consider questions about what it means for A to have trust, or faith, in B (trust as a binary relation).

Second, while this distinction will not play a role in the issues discussed in this paper, it should be noted that trust, a morally-inflected concept, is distinct from reliance, which does not carry the same moral weight. When one merely relies on a person, one takes a similar attitude towards her as one takes towards a clock—one does not feel that an inaccurate clock has betrayed one or failed to meet a responsibility (Baier 1984, 235). In contrast to relationships of reliance, trusting relationships carry moral weight because they possess the possibility of betrayal. The conceptual distinction between trust and reliance is not accurately reflected in our

common usage of the terms. We can, and do, use 'trust' to refer to both morally-inflected trust and reliance. It is this looser use of 'trust' that I will use in this paper. QWhile the distinction between trust and reliance is conceptually significant, questions about the moral weight of the cognitive attitude of counting on someone to do something will not play a role in this analysis. Thus, while I use the term 'trust', what I say about the cognitive attitude of trust also applies to the cognitive attitude involved in reliance.

3 Does trust require belief that the trusted will act as expected?

When I trust someone to ϕ, must I believe that she will ϕ? Several common and apparently contradictory observations about trust make it difficult to answer this question. Many authors (e.g., Jones 1996; McLeod 2002; Hieronymi 2008) share Baier's view that we cannot ordinarily choose to trust (Baier 1984, 244). This view that trust is involuntary is taken as evidence for the claim that trust requires belief that the trusted will act as expected, because it is commonly argued that belief is involuntary (Williams 1973). At first glance, it seems that trust is not under voluntary control. In some situations, if we do not trust someone, we cannot decide to change our position and trust her. If someone I know to be a sociopathic con artist invites me to trust her with my retirement savings, I cannot will myself to trust her to take care of my money. I could hand over my money to her, but I will not be able to do this with a trusting state of mind. Thus, the cognitive state of trust may appear to be non-voluntary belief.

However, this view that trust requires belief is undermined when we reflect on certain reasons we cite to explain our trust in others. Some of these reasons are not reasons one could cite to support a belief that someone will ϕ. There are at least three types of chosen trust that we engage in for reasons that do not support belief: therapeutic trust, coping trust, and corrective trust. Therapeutic trust has been examined in the trust literature, while coping and corrective

trust have been neglected.

One may choose to trust when one's trust can inspire positive change in the trusted party; this is *therapeutic trust*, which Karen Jones defines as "trust undertaken with the aim of bringing about trustworthiness" (Jones 2004, 5). Sometimes one might choose to trust in order to encourage, inspire, or motivate someone to live up to one's vision of the kind of person she could be (McGeer 2008). This choice to trust can be done despite the fact that one does not have sufficient evidence to support a belief that the trusted will act as expected. Richard Holton's shopkeeper example illustrates this:

> Suppose you run a small shop. And suppose you discover that the person you have recently employed has just been convicted of petty theft. Should you trust him with the till? It appears that you can really decide whether or not to do so. And again it appears that you can do so without believing that he is trustworthy. Perhaps you think trust is the best way to draw him back into the moral community. (Holton 1994, 63)

The shopkeeper is not trusting for reasons that constitute evidence for the belief that the thief will not steal. Similarly, a parent, knowing her teenager has a history of irresponsible behavior, may nonetheless choose to trust her teenager alone in the house in the hope that this will inspire the daughter to live up to the mother's parent's vision of the kind of responsible woman she can become.

A second type of chosen trust is *coping trust*. Sometimes we choose to plan based on the assumption that someone will do something because such planning helps us cope with the complexity of our interdependent lives and/or the anxiety that can accompany excessive attention to our vulnerability. It is a commonplace in the sociological and ethical literature on the professions that modern life involves being caught up in complex systems of expertise that we neither fully understand nor completely control (Giddens 1990, 2-3; Barber 1983; Pellegrino et al. 1991). While we may make some effort to base our trust in experts on

evidence of their trustworthiness, it is not uncommon to find oneself in a position where one needs to count upon some agent without having adequate evidence to support the belief that the agent is trustworthy. Nonetheless, one might still decide to trust because one determines that trust will simplify one's planning. Simplification can have psychological benefits, such as the reduction of anxiety, but it can also save time and other resources. Simplifying can reduce one's cognitive load, enabling one to pursue other valued projects. For example, consider a patient facing a minor medical procedure. She does not know her doctor well, and, due to recent media reports, knows that there is lax oversight of the medical community in her area. So the patient is in a state of doubt about whether she should trust her doctor. Foreseeing that she will only cause herself needless anxiety by pursuing a time-intensive investigation to gather evidence of her doctor's trustworthiness, she might choose to trust her doctor. This patient determines that she ought to engage in coping trust, a kind of trust which has received only scant attention in the literature.

Finally, recent work on testimonial injustice and implicit bias reveals the need for *corrective trust,* another previously unrecognized type of trust. A speaker sustains a testimonial injustice iff and only if "she receives a credibility deficit owing to identity prejudice in the hearer" (Fricker 2007, 28). When a hearer's prejudice about the speaker causes her to grant the speaker's testimony less credibility than would have been granted in the absence of the prejudice, the speaker suffers an injustice because she is undermined in her capacity as a knower (Fricker 2007, 44). Recent work on implicit bias shows that prejudice is not only reflected in one's explicitly held beliefs, but it can also operate unconsciously. Being a virtuous hearer involves more than cultivating an unprejudiced testimonial sensibility, so that one does not unjustly perceive the speaker as untrustworthy; it also includes correcting for both explicit and implicit bias that one recognizes has unjustly shaped one's perception of testimony:

> When the hearer suspects prejudice in her credibility judgement… she should shift intellectual gear out of spontaneous, unreflective mode and into active critical reflection in order to identify how far the suspected prejudice has influenced her judgement. If she finds that the low credibility judgement she has made of a speaker is due in part to prejudice, then she can correct this by revising the credibility upwards to compensate. (Fricker 2007, 91)

I argue that Fricker needs a notion of a type of chosen trust to account for this revision of credibility assessments. When a hearer suspects that a prejudice is undermining her ability to judge the trustworthiness of someone's testimony accurately, she ought, in many cases, to choose to trust the speaker by revising the credibility upwards to compensate for the prejudice. I call this *corrective trust*, and I argue that Fricker owes us an account of the cognitive attitude involved. One of the aims of this paper is to provide such an account.

Fricker's own analysis of epistemic injustice in *The Talented Mr Ripley* illustrates the need for corrective trust. In the film, Herbert Greenleaf arrives in Venice looking for his son, Dickie, who has recently disappeared. Dickie's girlfriend, Marge, rightly suspects that Dickie's friend, Ripley, is responsible for Dickie's disappearance. Greenleaf, however, has hired a private detective who tells Greenleaf that his son may have committed suicide. When Marge rejects this idea and tries to draw Greenleaf's attention to other evidence, Greenleaf dismisses her views saying, "Marge, there's female intuition, and then there are facts—" (qtd. in Fricker 2007, 88). Greenleaf unreflectively allows a prejudiced assessment of Marge's credibility to shape his views. What he ought to do is reflect on the prejudice behind his assessment and consciously choose to trust Marge. In choosing to trust, he would be a virtuous hearer and would avoid doing Marge a moral and epistemic harm.

These three types of chosen trust (therapeutic, coping, and corrective) create problems for the view that trust requires belief that the trusted will act as expected. Since belief is commonly thought to be an involuntary cognitive state, these are cases of a trusting state of

mind that is not belief. In addition, belief aims at truth, but agents often engage in therapeutic, coping, and corrective trust for non-evidential reasons. Agents engaged in therapeutic trust can have pragmatic, non-evidential, reasons to trust the party they hope to inspire to change. *A* might lack positive epistemic reasons to believe that *B* will φ, but *A* nonetheless has good reason to trust *B* to φ if she believes that trusting *B* to φ will have a positive effect on *B*'s character (Faulkner 2007, 314; McGeer 2008). Similarly, rational coping trust is grounded in a pragmatic reason—to simplify one's reasoning and avoid anxiety. We are limited beings, and often have to weigh the benefits of obtaining evidence of someone's trustworthiness against the costs of acquiring the evidence. When the costs of betrayal are low, it might not be worth the time or effort for *A* to gain adequate evidence for the proposition that *B* will φ, but *A* might see that trusting *B* to φ would simplify *A's* reasoning, reduce her cognitive load, and/or reduce her anxiety. Thus, *A* can be practically rational in making a cost/benefit assessment that gives *A* reason to trust *B* to φ without evidence that *B* will φ. Finally, another human limitation can practically rationalize corrective trust: human susceptibility to prejudices and false stereotypes, which corrupts our testimonial sensibility and causes us to commit testimonial injustices. Given *A*'s interest in avoiding doing harm to *B*, *A* can have a good reason to trust *B* to provide accurate testimony. Therefore, there can be good practical reasons for therapeutic, coping, and corrective trust, even when the truster lacks good epistemic reasons for trusting. Thus, reflecting on these examples of practically rational trust reveals that trust appears to involve a cognitive attitude which is, unlike belief, shaped by concerns other than truth.

One might object that, at least in the case of corrective trust, the reasons for trust are always epistemic, rather than moral. This is a worthwhile objection because often we do have epistemic reasons for corrective trust. If *A* is aware that a bias skews her evaluation of *B*'s

trustworthiness, then *A* can choose to engage in corrective trust in order to neutralize the effect of the bias and produce a more accurate appraisal of *B*'s testimony.   However, epistemic reasons do not exhaust the reasons why one would choose to trust to correct for a bias.   One argument for the existence of moral reasons for corrective trust starts from the premise that recognition of implicit bias involves recognition that one is cognitively impaired.   In such a state one may doubt that, even after recognizing the existence of a bias, one knows the extent to which the bias shapes one's trust in another's testimony.   Fricker often describes the virtuous hearer as revising upwards the degree of credibility she gives to the speaker in order to neutralize the effect of a bias (Fricker 2007, 91-2).   But a speaker hearer who recognizes that she is impaired may not know how severe the impairment is.   Thus, it may be hard to provide a solid epistemic justification for choosing any particular new degree of credibility to assign to the speaker's testimony.   In other words, recognition of an implicit bias may (and perhaps should) so radically undermine the hearer's trust in her testimonial sensibility that the epistemically rational thing to do is to accept that she is not a reliable judge of the truth of the speaker's testimony.   Imagine that Greenleaf has a dawning recognition that he is dismissing Marge due to his stereotype of the hysterical woman.   He might see that this harms her and wishes to avoid the harm, but, since this insight is part of a budding appreciation of sexism, he is not competent to judge how much credibility women like Marge deserve.   Despite lacking good epistemic reasons for granting Marge a degree of credibility, Greenleaf has a moral reason to trust Marge.   In such cases, epistemic reasons alone cannot provide reason to trust, while moral reasons can rationalize corrective trust.

4. A The unified account of trust: trust and practical reasoning*

In sumAs suggested by the considerations of the previous section, non-evidential

reasons for trusting and the voluntariness of some types of trust create problems for the view that trust requires belief that the trusted will act as expected. So what is the cognitive attitude constitutive of trust? Surely, it is sometimes belief. We often trust because the evidence that the trusted will act as expected compels us to trust, and in these situations it makes sense to say that the cognitive attitude is one of belief. What connection, if any, does this have to chosen, non-evidential trust? Intuitively, we should be able to provide an account of trust that unifies these two types of trust. They both involve an expectation that someone will do something, and, in both types of trust, that expectation makes us vulnerable. But what is the cognitive attitude that constitutes this expectation? In order to explain what corrective, therapeutic, and coping trust have in common with cases of trust that involve belief, we need an account of trust broader than those (e.g. Hieronymi 2008) which require that trust involves belief. On my account, A either *believes* or *accepts* that B will ϕ, and this belief or acceptance is the basis of A's practical reasoning. The following account accomplishes this: A trusts B to ϕ iff What makes this a unified account is that, in both types of trust, the proposition that B will ϕ is part of A's *adjusted cognitive background*. One's adjusted cognitive background includes all the propositions that one takes for granted in one's practical reasoning in a particular context. A proposition can be part of one's adjusted cognitive background through belief or acceptance. Thus on this account, A trusts B to ϕ iff A either *believes* or *accepts* that B will ϕ, and this belief or acceptance is the basis of A's practical reasoning. This unified account explains the role of both types of trust (trust-as-belief and trust-as-acceptance) in planning. When A trusts B to ϕ, A makes plans based on the assumption that B will ϕ. Having introduced this account of the cognitive attitude of trust in the context of the debate about the role of belief in trust, I now flesh out the account by explicating the key notions of belief, acceptance, and adjusted cognitive background.

4 The distinction between belief and acceptance

In "Practical Reasoning and Acceptance in a Context," Michael Bratman distinguishes between the mental states of belief and acceptance. In general, reasonable belief has four features that acceptance lacks (Bratman 1992, 3-4). First, reasonable belief is context-independent. My beliefs do not change as I move from one intellectual or practical context to another. Second, reasonable belief is "shaped primarily by evidence for what is believed and concern for the truth of what is believed" (Bratman 1992, 3). In other words, belief aims at truth. Third, we do not normally have direct voluntary control over our beliefs. Fourth, an agent's beliefs are subject to demands for consistency and coherence. In contrast, acceptance is context-dependent, shaped by factors other than evidence, voluntary, and exempt from demands for overall consistency across contexts.

Bratman's argument for the belief/acceptance distinction proceeds by presentation of several examples of context-dependent acceptance. The following example illustrates the four features of acceptance:

> The three of us need jointly to decide whether to build a house together. We agree to base our deliberations on the assumption that the total cost of the project will include the top of the estimated range offered by each of the sub-contractors. We facilitate our group deliberations and decisions by agreeing on a common framework of assumptions. We each accept these assumptions in this context, the context of our group's deliberations, even though it may well be that none of us believes these assumptions or accepts them in other, more individualistic contexts. (Bratman 1992, 7)

In this case, the building group has decided to use the highest estimated prices for materials and labor in its practical reasoning about the cost because it will simplify their work. This is a situation in which one can legitimately accept a set of assumptions in one context that one

would not accept in another context; for instance, if one were asked to place a bet on the cost of the house, one would not take the highest sub-contractor estimates for granted in one's calculations. Bratman's example also shows how reasonable acceptance, unlike belief, does not necessarily aim at truth. One can have pragmatic reasons for accepting a proposition in a given context; in this case, the group has a pragmatic interest in simplifying their deliberations. In addition, the group's acceptance of the cost framework is voluntary. Finally, we would find it strange were someone to criticize the group for accepting the high cost estimate on the grounds that it is inconsistent with the set of propositions the group accepted when they were trying to determine the cheapest price for the house. Sometimes we adopt a cognitive state of acceptance, which is not subject to the ideals of consistency and coherence across contexts.

An agent's beliefs form the "*default cognitive background*" that can be adjusted to suit practical reasoning about what to do in a specific context (Bratman 1992, 10). We bring to all contexts a set of involuntary beliefs that are subject to demands for evidence and global consistency. However, depending on the nature of the particular context at hand, we can bracket the belief that p, which is part of our cognitive background, or we can accept that p despite not maintaining the belief that p in the default background. We thus engage in practical reasoning in a specific context based on our "*context-relative adjusted cognitive background*" (Bratman 1992, 11). This adjusted cognitive background includes all the un-bracketed propositions we believe and all the propositions we have accepted for this particular context.

5 The cognitive attitude of trust

When A trusts B to ϕ, the proposition that B will ϕ is part of A's adjusted cognitive background. When we trust someone to do something, we make plans based on the

assumption that she will act as expected (Holton 1994; Pettit 1995). Sometimes we make plans based on the assumption that she will come through for us, even when we do not have good evidence to support the belief that she will. In these cases our trust is a matter of acceptance, rather than belief, that the trusted will act as expected. Consider therapeutic trust. Holton's shopkeeper may not have a context-independent belief that her employee will not steal from the till, but she may still choose to make plans for her business based on that assumption—she accepts the proposition that the employee will leave the money alone as a premise in her practical reasoning. Similarly, the patient engaged in coping trust decides to make plans for her post-procedure recovery based on the assumption that the doctor is worthy of her trust—she accepts the proposition that the doctor is trustworthy. Finally, Greenleaf might come to see that sexism clouds his ability to accurately assess Marge's credibility. In order to avoid the harm of testimonial injustice, he could choose to accept the proposition that Marge is a competent testifier as a premise in his practical reasoning. He could ask himself, "Assuming that she is right about Dickie, what kinds of questions should I ask Ripley, and where should I be searching for my son?" In this way, his corrective trust in Marge manifests itself in his accepting the fact of her competence in his planning for the search. Therefore, the first way that A can trust B to ϕ is by *accepting* the proposition that B will ϕ as part of A's adjusted cognitive background. It is this type of trust that, while often lacking epistemic support, is practically rational because accepting the proposition that B will ϕ can play a useful role in A's planning.

The other way the proposition that B will ϕ can be part of A's adjusted cognitive background is by A *believing* it. If A has the context-independent belief that B will ϕ, A may find herself in a specific context in which there is no reason to bracket this belief. In this case, A's trust in B to ϕ is a matter of A's belief that B will ϕ. Thus, this account is broader than accounts of trust which require that trust involves belief (e.g. Hieronymi 2008). Trust may

involve belief that someone will do something, but it may instead involve acceptance that someone will do something. In either case, trust involves taking the premise that someone will do something as the basis of one's practical reasoning about what to do, which is to say that trust involves having the proposition that someone will do something as part of one's adjusted cognitive background.

One of the virtues of this account is that it recognizes that trust is context-dependent. I may trust my friend not to tell my secret when exposure would do me little harm, while not trusting her with same secret in an environment where revelation would seriously damage my reputation. This context sensitivity is easily explained on this account. The difference between what we reasonably accept in one context and do not accept in another can be strongly influenced by asymmetries in the costs of error. Suppose I believe that my friend will keep my secret; nonetheless, I may choose to bracket this belief when the costs of error are high. Alternatively, I may doubt that she will keep my secret; nonetheless, I may choose to accept that she will keep it because the costs of error are low, and I think trusting her will have pragmatic benefits (for example, inspiring positive change in her). Our ability to adjust our cognitive background for practical reasoning in light of the details of the particular context of deliberation nicely explains why trust is context-dependent.

This account also explains why both evidential and non-evidential trust make one involve a cognitive attitude of vulnerabilitye. When one takes the premise that someone will do something as part of one's adjusted cognitive background, one may work it into one's plans that she will do it—the premise is available to use in one's practical reasoning. When one counts on someone in this way, one is vulnerable to having one's plans undermined. If the trusted fails to act as one expects, then the success of one's practical reasoning is threatened. One is vulnerable to planning and reasoning about what to do based on a false assumption.

This is the risk that one takes when one has the premise that someone will do something as part of one's adjusted cognitive background. Thus this account provides the unified account of both chosen and unchosen trust that is needed for the ethical and epistemic analysis of interdependence and vulnerability.

56 Objections and replies

56.1 Voluntary trust and epistemic constraints

What about the common view that trust is not subject to voluntary control? If some cases of trust involve acceptance, rather than belief, then it follows from this account that some cases of trust *are* under voluntary control. While the account denies that trusting can never be done at will, it can nonetheless accommodate some of the considerations behind the view that trust is involuntary. The key is to recognize that *acceptance* is not *supposition*. Supposition is used in hypothetical reasoning about what one *would* do were the supposition true. In contrast, reasoning on the basis of what one accepts is not hypothetical in this way. Supposition does not lead directly to action, but acceptance can (Bratman 1992, 9). For example, after watching a news story about people living with severe allergies, I begin to wonder how my life would change were I suddenly terribly allergic to my cat. I reason that if I were so afflicted, I would have to give up my cat, which might require me to find a new owner for my beloved pet. This is purely hypothetical reasoning. Since I know that I am not severely allergic now, this line of thought does not move me to call the animal shelter. In contrast, if one accepts that p, one takes oneself to have good reason *ceteris paribus* to act in light of p in that context. As I will show, because of the direct connection between acceptance and action, there are some epistemic constraints on acceptance. Recognition of these epistemic constraints can allay concerns about voluntary trust.

The reasons one has for accepting that p can be non-evidential, but that does not mean that evidence has no place to play in acceptance. There are some epistemic constraints on acceptance. One might wonder why there should be epistemic constraints on the premises we can use in our practical reasoning. It seems that we should be able to use any premise as part of our planning process. However, practical reasoning is reasoning about what *to do*. As such, it is distinct from daydreaming or hypothetical reasoning about what one would do in another distant possible world. In our practical reasoning we make plans about what to do in this world, and, therefore, our planning must be guided by premises that have some relevant semblance to the actual world. Thus, overwhelming amounts of evidence that p is not true (or not even approximately true) can make it unreasonable to accept p in virtually all contexts. A reasonable person could very well make hypothetical plans based on the supposition that she is allergic to her pet, but only an irrational person would non-hypothetically plan her life around that supposition in the face of overwhelming evidence to the contrary. Were we to ask a reasonable person to plan her life around it, she might tell us that she just cannot, no matter how hard she tries. I submit that the same can be true for trust.

In the vast majority of contexts, given an overwhelming amount of evidence against the trustworthiness of a person, we cannot trust her no matter how hard we try. The sociopathic con artist example illustrates this. I cannot choose to trust her to invest my money well. My sketch of the epistemic constraints on acceptance nicely accounts for this intuition; one cannot trust B to ϕ if one takes oneself to have overwhelming evidence that B will not ϕ, because one cannot accept that p if one takes oneself to have overwhelming evidence that p is not even approximately true. I cannot trust the sociopath to invest my money if I take myself to have overwhelming evidence that she will steal it. However, this epistemic constraint on trust only rules out trusting someone to do something one has overwhelming evidence to believe she will

not do; it does not rule out trusting someone when one is agnostic about whether she will act as expected. Holton makes this same point about reliance: "I do not need to have the belief that you will do what I rely [on] you to do, but I do need to lack the belief that you will fail" (Holton 1994, 71). In summary, unlike belief, acceptance can be voluntary, but that does not mean that we are free to accept whatever we wish, since acceptance is subject to stronger epistemic constraints than supposition is.

56.2 Full-fledged trust, pretensce, and acting-as-if

A second objection, proposed by Pamela Hieronymi, argues that trust without belief is merely a poor cousin of the type of trust of which we should give an account. She thinks that trusting without believing that the trusted will come through for you shows a lack of confidence in the trusted, and that "your lack of confidence betrays a lack of trust" (Hieronymi 2008, 6).

Hieronymi uses the following example to support this claim:
> Suppose that, in the morning, you and I agree to meet for dinner at a certain time at a certain restaurant to plan an upcoming event. Later in the day you learn that all my friends have decided to go to my favourite restaurant to celebrate a surprise promotion bestowed on one of them. You now doubt whether I will keep my engagement with you. You are not certain I will not, but then you are not certain I will either. You are in a state of doubt. In the face of your doubt, you decide to go to the restaurant and wait for me. (Hieronymi 2008, 6)

Hieronymi imagines that when I arrive at the restaurant, you tell me about your doubts and explain that you decided to go to the restaurant in spite of the doubts. She thinks that in this scenario, I will feel that your doubts express a lack of trust in me to keep our agreement (Hieronymi 2008, 6). Hieronymi calls trust accompanied by belief "full-fledged" trust. Full-fledged trust is a sort of ideal trust: "…even if one thinks the full-fledged sort of trust would be positively inappropriate in the circumstances, one can still imagine what it would be to have it,

and its inappropriateness is typically explained by features of the situation seen as regrettable" (Hieronymi 2008, 6-7). On this objection, insofar as my account allows room for trust based on acceptance rather than belief, it is an account of a less trusting sort of trust. What we really want is an account of the full-fledged, fully trusting sort of trust that requires belief.

In response, first, there is a reasonable alternative interpretation of Hieronymi's restaurant example. It seems to me that when you tell me about your doubts that I would arrive for dinner, it would be just as natural for me to respond, "I'm sorry I gave you reason to doubt me, but thanks for trusting me anyway." I do not think that we always take doubt and the absence of complete confidence to suggest a lack of trust. Second, even if we agree with Hieronymi that trust with belief is a more trusting sort of trust, I do not see why an account of trust should only focus on this narrow class of trust phenomena. Hieronymi acknowledges this response when she notes that some may dismiss her account of full-fledged trust as a mere 'purist's' notion of trust. But she argues that we ought to adopt the purist's notion of trust as "a natural refinement of our ordinary notion" because of the problems with the alternative accounts of trust (Hieronymi 2008, 2).

One such problem stems from the sensible demand that trust be distinguished from mere pretense of trust. Many authors worry that trusting without belief is too similar to acting as if one trusts without actually trusting (Baker 1987; Holton 1994; Hardin 2002; Hieronymi 2008). This worry is thought to be particularly pressing when it comes to trusting others' testimony. Suppose that my friend tells me that she is innocent of the crime of which she has been accused. It might seem right to say that what my friend wants of me is to believe that she is innocent, and she might charge me with failing to trust her if I do not believe in her innocence. Judith Baker uses this scenario to make the following argument that trust must require belief:

> Someone might try to distinguish trust from genuine or full belief. Trust, on such a view, would be a watered down variant of belief, something more like pretence or acting-as-if something were true. But this is to view trust as a non-serious form of belief. Whereas what one demands from one's friends is belief, not pretence, that one is innocent. (Baker 1987, 6)

Now I agree that when we trust a friend we do not merely act as if we believe her, but Baker presents us with a false dichotomy between believing and acting as if one believes. Acceptance is a cognitive state that is neither belief nor pretense. The reasoning involved in pretense has a more indirect and complex connection with action than does simply reasoning on the basis of a premise that one accepts (Bratman 1992, 9). Recall my friend who has been accused of a crime of which she claims innocence. It may be reasonable for me, being a loyal friend, to not use the accusation against her as a reason not to invite her to my upcoming dinner party (Bratman 1992, 8). So I trust her and simply plan my guest list on the assumption that she is innocent.Consider an example Bratman attributes to Michael Dummett:

> My close friend has been accused of a terrible crime, the evidence of his guilt is strong, but my friend insists on his innocence. Despite the evidence of guilt, my close friendship may argue for assuming, in my ordinary practical reasoning and action, that he is innocent of the charge. In making plans for a dinner party, for example, such considerations of loyalty might make it reasonable for me to take his innocence for granted and so not use this issue to preclude inviting him. (Bratman 1992, 8)

The practical reasoning involved in the decision to invite the friend to dinner is relatively simple. In contrast, the reasoning involved in merely pretending to trust the friend's profession of innocence must involve considerations about how to keep up the appearance of trust (e.g., planning how to give the friend the impression that everything is normal, and planning how to

cover up any hints of my pretense).   Therefore, the planning involved in instead of just reasoning as one normally would and inviting the friend, one decides to act as usual because otherwise the friend might detect the pretense—trust-as-acceptance is simpler than the reasoning involved in pretending to trust. This example illustrates how acceptance of testimony need not be mere pretense. In this way, the worry that my account of trust is incapable of marking the distinction between trusting and pretending to trust is alleviated.

It is on this question of the distinction between trusting and acting as if one trusts that my account differs most from its closest relative—Holton's account of trust as reliance from the participant stance. Holton agrees with my view that trust does not require belief and that we can sometimes choose to trust. When I trust or rely on someone to do something, "I plan on the supposition that they will do it" (Holton 1994, 72). This is very similar to saying that I make the assumption that she will do it part of my adjusted cognitive background.   Despite these similarities, Holton does not clearly mark the distinction between trusting and acting as if one trusts. He gives the example of trust circle games to illustrate his account of choosing to trust. In these games, the participant stands in a circle of people who are supposed to catch her as she falls backwards. The participant closes her eyes and lets herself fall backwards into the circle. Holton says that at the moment before one falls, one can choose to fall despite having some doubts about whether one will be caught. To explain what happens when one makes that choice, Holton draws the following analogy: "Just as the non-believer in the [religiously] strict society can decide to act as a believer would, so I can decide to act on the supposition that you will catch me. That is to decide to rely on you" (Holton 1994, 69). Deciding to trust the people in the trust circle is compared here to acting as if one believes the religious doctrines of a strict religious society.   In fact, Holton explicitly characterizes trust as "a kind of acting-as-if" (Holton 1994, 73).

The problem with Holton's account is that trusting and acting as if one trusts are different. One can act as if one takes p as a premise in one's practical reasoning without actually doing so. Both parts of Holton's analogy illustrate this. The non-believer who acts as a believer may do so simply to avoid shunning without taking any of the believer's religious doctrines as premises in her practical reasoning. Similarly, I can act as if I trust my friend, by telling her a secret, without in any way taking the proposition that she will keep it as a premise in my practical reasoning. I may pretend to trust her in order to catch her in the act of spreading my secrets.   The problem with Holton's account is not that he says trust involves acting onplanning on the supposition that the trusted will act as expected. The problem is that Holton maintains that acting onplanning on such a supposition is a case of acting-as-if. My account does not conflate these two notions.

It is crucial for an account of trust to distinguish between acting-as-if and trusting, because acting as if one trusts will undermine some of the reasons we have for trust. This is particularly true of corrective, therapeutic, and coping trust. First, one of the reasons to engage in corrective trust is to avoid committing the testimonial injustice of disrespecting a speaker as a knower. Merely acting as if a speaker is a trustworthy testifier does not show respect for this aspect of the speaker's humanity; it is merely pretending to respect her as a knower. Second, the reason to engage in therapeutic trust is to inspire the trusted to change by holding out to them a vision of the kind of person one thinks her capable of being. However, merely acting as if one trusts another is an "unstable mechanism" for generating the benefits of therapeutic trust: if B detects that A does not in fact hold the favorable vision of her, then B "will lose the incentive provided by [A]'s trust to act in a trust-responsive way" (McGeer 2008, 16). Thus, acting as if one trusts can undermine the very purpose of therapeutic trust. Planning on the basis of the acceptance that B will ϕ, however, does show B that one holds the positive vision

of her. Third, acting as if one trusts undermines coping trust, the purpose of which is to simplify one's planning. Consider the role of coping trust in anxiety reduction. Acting as if one trusts may require one to dwell on numerous potentially anxiety-provoking scenarios, e.g. the possibility that one's pretense of trust will be detected. In addition, if A's practical reasoning is not based on the assumption that B will ϕ, there is no reason for A not to dwell on the possibility of being let down by B, and this is the main anxiety-reducing reason for choosing to engage in coping trust. Therefore, in order to account for these three types of chosen trust, we need an account of trust that does not conflate acting-as-if and trusting.

56.3 Back-up planning

Finally, one might object that my account cannot explain the role of back-up planning in rational trust. Suppose that A believes that B will ϕ and acts in ways that make her vulnerable were B to fail to ϕ. However, A also makes back-up plans for what to do if B fails to ϕ. The objector claims that intuitively we want to say that A does trust in this situation; A is just a careful, rational truster, one who trusts but plays it safe by devising a contingency plan. However, my account, which locates trust in planning, seems to deny that A trusts in this case, because A plans on the basis of the assumption that B will not ϕ. Thus, this appears to be a counterexample to the account.

However, back-up planning is more complex than the objection allows. Two types of back-up planning must be considered: hypothetical and action-oriented. Hypothetical back-up plans involve the attitude of supposition discussed earlier. In this kind of back-up planning, one imagines what one would do were something to be the case, but this planning lacks a direct link to action. Just as I might imagine what I would do were I to develop an allergy to my cat, A might imagine what she would do were B to fail to ϕ. Suppose that this is the kind of back-

up planning at issue in the proposed counterexample. The additional hypothetical back-up plans do not affect A's trust in B. As long as A believes that B will ϕ, makes plans on the basis of this belief, and takes herself to have good reason *ceteris paribus* to act in light of this belief in that context, then my account implies that A trusts B to ϕ. Any additional hypothetical back-up plans do not undermine A's trust. This is contrary to the objector's claim that my account says A does not trust B to ϕ in this case. So if this is the kind of planning involved in the example, it is not a genuine counterexample.

But suppose the example involves another kind of back-up planning which is directly tied to action. Sometimes we make back-up plans from an action-oriented stance. For example, imagine I have scheduled a guest speaker for my class. I trust my guest speaker to show up, and I make plans for how I will introduce her to my students. These plans spur me to action (e.g., I write notes for myself of her accomplishments, I take the notes with me, etc.). But I also reason through what I will do if she forgets the appointment, and I make plans for how I will teach the class in her absence. These are not merely hypothetical plans, they spur me to action (e.g., I look for a film to show, I make notes about questions to ask about the film, and I take the film with me, etc.). Consider my attitude when I am making these plans before I have started to write my notes or look for a film. Is it a trusting attitude? Intuitively, we want to say that it is not, or that, at best, it is a mixed case of trust and distrust. When I think to myself, "If she does not come, I'll have 55 minutes of class time to fill. So I should devise something to teach," I am searching for ways to reduce my vulnerability. We take this kind of back-up planning as a sign of distrust. This is because trust involves an attitude of "acceptance of some degree of vulnerability to another's power, in the confidence that this power will not be used to harm or hurt one" (Baier 2007, 136). When one is confident that the trusted will act as expected, "one forgoes searching (at the time) for ways to reduce such vulnerability" (Jones

2004, 8).   To the extent that *A* uses back-up planning to search for ways to reduce her vulnerability, she lacks the kind of confidence reflected in taking the premise that *B* will ϕ as the basis of her practical reasoning. If *A* engages in enough of this kind of back-up planning, then I think we want to say that she does not trust *B* to ϕ.   So this example does not pose a counterexample to the account, because we do not have the intuition that it is a case of trust. If *A* engages in a small amount of this kind of back-up planning, I think we want to say that *A* has a mixed attitude towards *B*: she both trusts and does not trust *B*, which can also be explained on my account.   In such a condition, the rational thing to do is to engage in back-up planning which reflects our lack of confidence in the truster. *A* may believe that *B* will ϕ, but the belief is weak and *A* takes herself to have sufficiently strong practical reasons to bracket the belief in contexts where vulnerability reduction is salient. Therefore, locating trust in our resources for planning also helps account for the ways in which rational agents attempt to reduce their vulnerability by making back-up plans.

67 Conclusion

Failure to account for the role of trust in planning has led many authors to dismiss choosing to trust as either impossible or irrational. However, agents often have good reason to make plans based on the assumption that someone will do something. Even in the absence of evidence that *B* will ϕ, *A* can have therapeutic, coping, or corrective reasons to trust.   To accommodate such types of trust requires an account of the cognitive attitude of trust broader than belief. Trust involves taking the proposition that someone will do something as a premise in one's practical reasoning, which can be a matter of believing or accepting the proposition. Trusting acceptance is neither a watered down kind of belief, nor an act of pretensce. Instead, it is a rationally constrained cognitive attitude that can be adopted at will for the purpose of

planning.

References

Baier, A. (1984). Trust and Antitrust. *Ethics*, 96(2), 231-260.

Baier, A. (1994). *Moral Prejudices*. Cambridge, MA: Harvard University Press.

Baier, A. (2007). Trust, Suffering, and the Aesculapian Virtues. In R. L. Walker & P. J. Ivanhoe (Eds.), *Working Virtue: Virtue Ethics and Contemporary Moral Problems*. New York: Oxford University Press.

Baker, J. (1987). Trust and Rationality. *Pacific Philosophical Quarterly*, 68(10), 1-13.

Barber, B. (1983). *The Logic and Limits of Trust*. New Brunswick, NJ: Rutgers University Press.

Bratman, M. (1992). Practical Reasoning and Acceptance in a Context. *Mind*, 101(401), 1-15.

Cohen, L.J. (1989). Belief and Acceptance. *Mind*, 98(391), 367-89.

Cohen, L.J. (1992). *An Essay on Belief and Acceptance*. Oxford: Clarendon Press.

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

Engel, P. (Ed.) (2000). *Believing and Accepting*. Boston: Kluwer Academic Publishers.

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

Faulkner, P. (2011). *Knowledge on Trust*. New York: Oxford University Press.

Fricker, M. (2007). *Epistemic Injustice: Power & the Ethics of Knowing*. New York: Oxford University Press.

Giddens, A. (1990). *The Consequences of Modernity*. Stanford, CA: Stanford University Press.

Ginet, C. (2001). Deciding to Believe. In M. Steup (Ed.), *Knowledge, Truth, and Duty*. New York: Oxford University Press.

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

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

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

Hieronymi, P. (2008). The Reasons of Trust. *Australasian Journal of Philosophy*, 86(2), 1-24.

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. (2002). The Politics of Credibility. In. L. M. Antony & C. E. Witt (Eds.), *A Mind of One's Own: Feminist Essays on Reason and Objectivity* (2nd Ed.). Boulder, CO: Westview Press.

Jones, K. (2004). Trust and Terror. In P. DesAutels & M. U. Walker (Eds.), *Moral Psychology: Feminist Ethics and Social Theory*. New York: Rowman & Littlefield.

Kelly, D. & Roeddert, E. (2008). Racial Cognition and the Ethics of Implicit Bias. *Philosophy Compass*, 3(3), 522-540.

Lehrer, K. (1979). The Gettier Problem and the Analysis of Knowledge. In G. S. Pappas (Ed.), *Justification and Knowledge*. Dordrecht: Reidel.

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

McGeer, V. (2008). Trust, Hope and Empowerment. *Australasian Journal of Philosophy*, 86 (2), 1-18.

Nagel, J. (2010). Epistemic Anxiety and Adaptive Invariantism. *Philosophical Perspectives*, 24 (1), 407-435.

Perry, J. (1980). Belief and Acceptance. *Midwest Studies in Philosophy*, 5(1), 533-542.

Pellegrino, E., Veatch, R. & Langan, J. (Eds.) (1991). *Ethics, Trust, and The Professions*.

Washington, DC: Georgetown University Press.

Pettit, P. (1995). The Cunning of Trust. *Philosophy & Public Affairs*, 24(3), 202-225.

Potter, N. (2002). *How Can I Be Trusted? A Virtue Theory of Trustworthiness*. New York: Rowman & Littlefield.

Project Implicit Information Website. (n.d.) Retrieved January 23, 2012, from http://www.projectimplicit.net/index.php.

Scheman, N. (2001). Epistemology Resuscitated: Objectivity as Trustworthiness. In N. Tuana & S. Morgen (Eds.), *Engendering Rationalities*. Albany, NY: State University of New York Press, Albany.

van Fraassen, B. (1980). *The Scientific Image*. Oxford: Clarendon Press.

Velleman, J. D. (2000). *The Possibility of Practical Reason*. Oxford: Clarendon Press.

Walker M. U. (2006). *Moral Repair: Reconstructing Moral Relations after Wrongdoing*. New York: Cambridge University Press.

Williams, B. (1973). Deciding to Believe. In *Problems of the Self*. New York: Cambridge University Press.