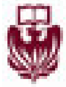

CHICAGO JOURNALS

Trust as an Affective Attitude
Author(s): Karen Jones
Source: *Ethics*, Vol. 107, No. 1 (Oct., 1996), pp. 4-25
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/2382241
Accessed: 08/07/2009 14:33

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=ucpress.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.

The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *Ethics*.

http://www.jstor.org

Trust as an Affective Attitude*

Karen Jones

I. INTRODUCTION

In this article I defend an account of trust according to which trust is an attitude of optimism that the goodwill and competence of another will extend to cover the domain of our interaction with her, together with the expectation that the one trusted will be directly and favorably moved by the thought that we are counting on her. The attitude of optimism is to be cashed out not primarily in terms of beliefs about the other's trustworthiness, but rather—in accordance with certain contemporary accounts of the emotions[1]—in terms of a distinctive, and affectively loaded, way of seeing the one trusted. This way of seeing the other, with its constitutive patterns of attention and tendencies of interpretation, explains the willingness of trusters to let those trusted get dangerously near the things they care about. This account is presented and defended in the first two sections of the article.

Any account of what trust is sets constraints on what can be said about the justification conditions of trust. Thus, if a theorist analyzes trust as (perhaps among other things) a belief that the one trusted will have and display goodwill toward the one who trusts, then that

* I would like to thank the editors of *Ethics,* Judith Baker, and an anonymous referee for comments that greatly improved this article and for suggesting a new title. Lively discussions with Bennett Helm, Martha Nussbaum, Naomi Scheman, and audiences at Cornell University and UCLA helped considerably. Special thanks are owed to Allen Wood for a discussion of evidentialism and to Terry Irwin for helpful comments on numerous drafts.

1. Amélie Oksenberg Rorty, "Explaining Emotions," in *Explaining Emotions,* ed. Amélie Oksenberg Rorty (Berkeley: University of California Press, 1980); Cheshire Calhoun, "Cognitive Emotions?" in *What Is an Emotion?* ed. Cheshire Calhoun and Robert C. Solomon (New York: Oxford University Press, 1984), pp. 327–42; Ronald de Sousa, "The Rationality of Emotions" (originally published 1979), in Rorty, ed., pp. 127–52, and *The Rationality of Emotion* (Cambridge, Mass.: MIT Press, 1987).

Ethics 107 (October 1996): 4–25
© 1996 by The University of Chicago. All rights reserved. 0014-1704/97/0701-0001$01.00

theorist has committed herself to saying that trust is justified only if the one who trusts is justified in forming the belief constitutive of trust. In the fourth section of the article, I take up the question of the justification conditions of trust. While a full account of the justification conditions of trust is beyond the scope of this article, I identify the key variables affecting the justifiedness of trust. An account of trust that makes affect central has an unexpected payoff: it is able to view a wide range of our trustings—including many of those undertaken for instrumental reasons—as justified. Moreover, it is able to do this without taking a stance on evidentialism, or the doctrine that we should not believe anything without sufficient evidence. Since trust is not primarily a belief, it falls outside the scope of the evidentialist thesis.

It is necessary, first, to get clearer about the target of my investigation. The word 'trust' is used in a variety of expressions, ranging from "Trust you to do something like that!" to "We trust you have enjoyed your flight with Air New Zealand," to "Othello's trust in Iago was misplaced." In the first sentence 'trust' is used ironically, although it brings with it from its nonironic uses the idea of expectations having been met, while in the second it politely conveys something intermediate between an expectation and a hope. Sometimes the word 'trust' is used to convey any sort of delegated responsibility, especially one where checking up is difficult or precluded. Thus a politician of such egregious ethical turpitude that she has long since ceased to be trusted by any of her constituents can nonetheless, on the exposure of some new failing, be said to have once again violated the public's trust. My task is thus not to explicate the meaning of the word 'trust' wherever it occurs, since there is no one common phenomenon that all uses of the word 'trust' pick out. Instead, my target is the sense conveyed in our third example: "Othello's trust in Iago was misplaced." That is to say, my target is interpersonal trust. But this is not a narrow target: it is the trust always found in friendship, often found between professionals and their clients, sometimes found between strangers, and sometimes, even, between people and their governments. My task is therefore an explanatory task, the success of which is to be tested by how well it lets us understand this everyday phenomenon—how well, that is, it can account for the similarities and differences between interpersonal trust relations of the sorts just listed.

II. AN ACCOUNT OF TRUST

A. The Basic Model

Trusting is composed of two elements, one cognitive and one affective or emotional. (I say "affective" rather than "noncognitive" because affective states can themselves contain a significant cognitive component.) Roughly, to trust someone is to have an attitude of optimism about her goodwill and to have the confident expectation that, when

the need arises, the one trusted will be directly and favorably moved by the thought that you are counting on her. If A's attitude toward B (in a given domain of interaction) is predominantly characterized by optimism about B's goodwill and by the expectation that B will be directly and favorably moved by the thought that A is counting on her, then A has a trusting relationship with B (within that domain). There can be moments of trust within relationships that are not, in general, characterized by trust, although if the attitude and expectation are too fleeting, it would not be correct to say that A trusts B. The attitude and expectation characteristic of trust combine to explain why trusters are willing, when the need arises, to rely on those they trust.[2]

In the standard case, the confident expectation that the one trusted will respond directly and favorably to the thought that the truster is counting on them is itself grounded in the attitude of optimism; thus the attitude of optimism is central. This account of the two aspects of trust requires further elaboration and refinement. It also needs to be shown why we should think that both are necessary for trust and why we should think that together they amount to a satisfying account.

First, though, I should explain what I mean by 'optimism', for the word has connotations that are apt to be misleading. The attitude of optimism is directed at the goodwill of another. I can trust someone with whom I'm engaged in a very difficult endeavor even though I have no optimism about the success of our joint task; thus, trust does not involve a general tendency to look on the bright side. However, trust does lead one to anticipate that the other will have and display goodwill, and this is the aspect of optimism that I want to highlight—the way optimism leads us to anticipate a favorable outcome. Throughout the article, though, I do not want 'optimism' to suggest a general tendency to look on the bright side. With that in mind, we can turn to the task of refining our characterization of trust.

At the center of trust is an attitude of optimism about the other person's goodwill. But optimism about goodwill is not sufficient, for some people have very good wills but very little competence, and the

2. The account I develop here is indebted to Annette Baier's account—in "Trust and Antitrust," *Ethics* 96 (1986): 231–60, and "The Pathologies of Trust" and "Appropriate Trust," delivered at Princeton University as the Tanner Lectures on Human Values (*Tanner Lectures on Human Values,* vol. 13 [Salt Lake City: University of Utah Press, 1992])—most significantly in the following ways: (i) in maintaining a distinction between trust and reliance, (ii) in acknowledging the importance of the competence of the other (Baier, "Trust and Antitrust," p. 239, and *Tanner Lectures on Human Values,* pp. 111–12), and (iii) in recognizing that trust can be faked (*Tanner Lectures on Human Values,* p. 112). The difference in our positions will become clear in Sec. III.

incompetent deserve our trust almost as little as the malicious. (Almost, but not quite, for the incompetent might sometimes get things right, whereas the malicious will get things right only to the extent that they are incompetent.) Thus, we should say that trust is optimism about the goodwill and *competence* of another. The position requires additional refinement: except perhaps with our most trusted intimates, the optimism we bear is seldom global. This is not to say that the optimism itself is qualified and instead of being unreserved optimism is a qualified or restricted optimism. What is qualified is not the optimism itself, but the domain over which it extends. So, for example, the optimism we have about the goodwill and competence of strangers does not extend very far. We expect their goodwill to extend to not harming us as we go about our business and their competence to consist in an understanding of the norms for interaction between strangers. For a man to run up at full speed behind a woman on an ill-lit street is to display a lack of such competence, and, even if he was simply out for a late night run and meant no harm, he has given the woman reason to distrust him. When we trust professionals, from plumbers to physicians, we expect of them a technical competence (and minimal decency). However, the competence we expect in trusting need not be technical: when we trust a friend, the competence we expect them to display is a kind of *moral* competence. We expect a friend to understand loyalty, kindness, and generosity, and what they call for in various situations.

There are a number of reasons why we might think that a person will have and display goodwill in the domain of our interaction with her. Perhaps she harbors friendly feelings toward us; in that case, the goodwill is grounded on personal liking. Or perhaps she is generally benevolent, or honest, or conscientious, and so on. The formulation is meant to be neutral between these reasons for thinking that a person's goodwill extends to cover the domain of our interaction.

It might be thought odd to claim that trust centrally involves an affective *attitude,* but this analysis is borne out by considering distrust. Distrust is trust's contrary and is synonymous with wary suspicion. Distrust is pessimism about the goodwill and competence of another (again, relativized to a certain domain), but to be pessimistic about someone's goodwill is to expect that it is likely that she will harm your interests, and thus to treat her warily and with suspicion.

The analysis is further borne out by considering a parallel between trust and self-confidence. It seems intuitively correct to say that self-confidence involves an affective attitude. To have self-confidence is to be optimistic about one's competence (in the domain in question) and to have the expectation that one will be able to bring about a favorable outcome. Sometimes we use the phrase 'trust yourself' as roughly interchangeable with 'be self-confident'. There seems, though, to be

an important difference between the two: with self-confidence, and its lack, self-doubt, we are worried about our capacity, rather than our will. 'Trust yourself' has application precisely because parts of ourselves can sometimes stand in the kind of external relation to other parts that makes their interaction more like the interaction between two persons. We need to trust ourselves when we are worried about the possibility of self-sabotage, about the possibility that some not fully conscious part of ourselves might be operating from motives other than our professed ones.

While trust essentially involves an attitude of optimism that the goodwill and competence of another will extend to the domain of our interaction with her, it is not exhausted by such an attitude. The affective element of trust needs to be supplemented with an expectation, namely, the expectation that the one trusted will be directly and favorably moved by the thought that someone is counting on her. Being directly and favorably moved by this thought may not give the one trusted an overriding motive; acting on such a thought, could, for example, be tempered by other concerns or by thinking that what the one who trusts is counting on is not, under the circumstances, in her own best interests. Nevertheless, one is not trustworthy unless one is willing to give significant weight to the fact that the other is counting on one, and so will not let that consideration be overruled by just any other concern one has. For this reason, one would not trust if one thought that the fact that one was counting on someone, while always being taken into account, would nonetheless be reliably overridden by other considerations. Were that the case, then, from the point of view of the truster, the other would appear unwilling to give enough weight to the thought that she was counting on her. If someone thought another would give this much weight and not more to the thought that she was being counted on, then she would not willingly rely on her if the need to do so were to arise. However, the truster's expectation need not amount to an expectation of actual performance in every case. Someone doesn't show herself untrustworthy simply because there are occasions on which the thought that someone is counting on her is not a consideration that she can let prevail. Further, when the attitude and expectation lead the truster to willingly rely on the one trusted, there may be (though there does not have to be) some vagueness about what it is the truster is counting on her for. There may be a number of ways of adequately responding to the thought that you are being counted on, which is why trusting is associated with discretionary powers.[3]

3. That trust involves discretionary powers is first noted by Baier ("Trust and Antitrust," pp. 236–40). However, I think she rather overstates the case in claiming that trust always involves discretionary powers.

The qualification 'directly' in "directly and favorably moved by the thought that someone is counting on her" is required to distinguish trusting from certain cases of mere reliance. I might, for example, know that you will be moved by the thought that I am counting on you because you fear my retaliation if you let me down. If I believe that you will be directly moved by fear and only indirectly moved by the thought that I am counting on you, our relationship is not, on my analysis, one of trust.

There are two ways to see that an expectation has to be added to the affective component of trust in order to have an adequate account: by considering unwelcome trust and by considering the ways in which a reliably benevolent person's actions and motives might yet fall short of the actions and motives that we would demand of someone we trust.

We do not always welcome trust. Sometimes someone's trust in us can feel coercive. When it does, we don't usually complain about the person's having an attitude of optimism about our goodwill and competence, or even about her displaying such optimism in her interaction toward us, for it is rare that we would find such an attitude unwelcome. (Although there can be such cases: as when, for example, we find it impertinent that someone has attributed goodwill to us with respect to a particular domain of interaction.) In the standard case, however, what we object to when we do not welcome someone's trust is that, in giving it, she expects that we will be directly moved by the thought that she is counting on us and, for one reason or another, we do not want to have to take such expectations into account, across the range of interactions the truster wants. (If we are morally decent, we do not find the trust common between strangers coercive, because what is demanded of us is minimal.) We would rather that the one trusting did not expect us to respond to her counting on us because we would rather not have her count on us. We may, for example, feel that we cannot live up to her expectations, or we may have reservations about what such expectations will amount to in a given case, or we may feel that too many people are already counting on us and that one more is a burden we would rather not have.

Perhaps not everyone will be convinced by this argument. It might be thought that we never object to someone's trust, as such, but only to their *entrusting* certain things to us. When someone entrusts something to our care they expect us to respond to the fact that they are counting on us. Cases of objectionable or unwelcome trust are always cases of unwelcome entrusting, as when, for example, you burden me with your secrets. Thus, having the expectation that another will be directly moved by your counting on them is part of entrusting but, for all that's been said so far, not part of trusting. (When I discuss Baier's account of trust, I'll return to the connection between trust and entrusting in more detail.) However, we can see that this objection

cannot be right, for I can find your trust a burden even when you have not entrusted anything in particular to me. Moreover, since we can entrust where we don't trust (I might know, for example, that you will take good care of whatever I entrust to you because you wouldn't dare do otherwise), it seems that if entrusting *sometimes* involves the expectation that the other will be directly moved by the thought that we are counting on them, then that expectation must be part of our trusting rather than our entrusting.

The second consideration in favor of supposing that trust must involve an expectation as well as an attitude is that someone who isn't at all directly moved by the thought that you're counting on them but is, let us suppose, reliably benevolent toward you, is reliable rather than trustworthy. Suppose that the only operative motive in your interaction with me is concern about my well-being. Regardless of what I count on you to do, you do it only if it maximizes my well-being, and if it does that, you would do it anyway, whether or not I counted on you to do so. I would be justified in having an attitude of optimism about your goodwill while refraining from seeing you as trustworthy.

It might be objected that it is only in cases where optimism about goodwill is grounded in perceived benevolence that we need also attribute to the truster the expectation that the other will be directly and favorably moved by the thought that she is being counted on. To demand in it all cases, or even in the majority of cases, is to make one's analysis overly narrow and vulnerable to counterexample. But this seems to me mistaken. Consider trust in physicians.[4] The objection asks, "Isn't it enough for me to count as trusting my physician if I view her as a person of integrity and competence who cares about the interests of her patients? Why must I also expect that she will be responsive to my counting on her?" The answer is that we hope that what the physician takes to constitute acting with integrity and takes to constitute the interests of her patients will be, at least in part, shaped by the expectations of those patients. And if a physician refuses to allow the expectations of her patients to shape her understanding of what, here and now, good medical practice consists in, her patients would not be justified in trusting her. (This explains why a physician might have reservations about having someone as her patient: if she feels that she will have objections to living up to her patient's expectations, she will think it difficult to maintain the proper relationship of trust.) For this reason, it would be a mistake to think that the ideally moral are always properly trusted. While it might be true that the ideally moral are properly trusted by those who are themselves ideally

4. Thanks to the editors of *Ethics* for this example.

moral, it doesn't follow that they are properly trusted by those who are not.

I have claimed that trust is composed of two elements: an affective attitude of optimism about the goodwill and competence of another as it extends to the domain of our interaction and, further, an expectation that the one trusted will be directly and favorably moved by the thought that you are counting on them. Our expectation is, in the typical case, grounded in the attitude of optimism. That is to say, we expect that the other will react favorably to our counting on them *because* we are optimistic about their goodwill. Our expectation is usually grounded in the very same evidence that grounds our attitude of optimism. Thus the attitude of optimism is central.

B. The Attitude of Optimism

We now have a sketch of an account of trust. But it remains a sketch insofar as we do not yet have a firm grip on what is meant by saying that trust is, among other things, an affective attitude of optimism about the goodwill of another.

According to one influential account of the emotions, held in various forms by Rorty, Calhoun, and de Sousa,[5] emotions are partly constituted by patterns of salience and tendencies of interpretation. An emotion suggests a particular line of inquiry and makes some beliefs seem compelling and others not, on account of the way the emotion gets us to focus on a partial field of evidence. Emotions are thus not primarily beliefs, although they do tend to give rise to beliefs; instead they are distinctive ways of seeing a situation. In resentment, for example, the object of resentment might be seen as a "manipulative exploiter."[6] Similarly, the claim being advanced here is that the attitude of optimism constitutive of trust is a distinctive way of seeing another. This way of seeing the other is constituted by a distinctive trusting cognitive set, which makes one's willingness to rely on the other seem reasonable.

The cognitive set constitutive of trust restricts the interpretations of another's behavior and motives that we consider. It also restricts the interpretations we will consider as possibly applying to situations and the kinds of inferences we will make about the likely actions of another. Consider the following exchange:

> Iago: My lord, you know I love you
> Othello: I think thou dost;
> And, for I know thou'rt full of love and honesty

5. Rorty, "Explaining Emotions"; Calhoun; and de Sousa, "The Rationality of Emotions" and *The Rationality of Emotion.*

6. For a discussion of the notion of cognitive sets and of the cognitive set involved in resentment in particular, see Calhoun.

> And weigh'st thy words before thou givest them breath,
> Therefore these stops of thine fright me the more;
> For such things in a false disloyal knave
> Are tricks of custom; but in a man that's just
> They're close dilations, working from the heart
> That passion cannot rule. (*Othello*, 3.3.117–23)

Othello trusts Iago and interprets his words and behavior in the light of his trust. Had Othello not trusted Iago, he would have been able to see Iago's speech and the very fact of his interference for what they were, malicious attempts to harm him. Trust restricts the interpretations we will consider as possibly applying to the words and actions of another. When we can—and sometimes even if doing so requires ingenuity—we will give such words and actions a favorable interpretation as consistent with the goodwill of the other. Trusting thus functions analogously to blinkered vision: it shields from view a whole range of interpretations about the motives of another and restricts the inferences we will make about the likely actions of another. Trusting thus opens one up to harm, for it gives rise to selective interpretation, which means that one can be fooled, that the truth might lie, as it were, outside one's gaze. Because we impute honorable motives to those we trust, and typically do not even stop to consider the harms they might cause if they have dishonorable motives, we are willing to rely on those we trust. The harms they might cause through failure of goodwill are not in view because the possibility that their will is other than good is not in view. What in the absence of trust would be taken to be a reason for jealousy, for wary suspicion, or for action to protect my interests will not be so taken when there is trust.[7]

It is because the one trusted is viewed through the affective lens of trust that those who trust are—usually cheerfully, and often on the basis of the smallest evidence—willing to risk depending on the one trusted. Someone might object that it is possible to have this distinctive way of seeing another without trusting her. You might see her in this way, but resist the appearance, and struggle to keep nontrusting interpretations in mind. We can see the force of this objection by considering a possible parallel with phobic emotions. If I have a phobic fear of spiders, I still fear them, even though, let us suppose, once I'm aware of my fear I make every effort to resist the patterns of salience and tendencies of interpretation that constitute fear.[8] If trust, on account of having an affective component, has features in common with emotions, it should be the case that there can be "phobic" trusting,

7. Ronald de Sousa discusses *Othello* in the context of how control of perceptual focus can give rise to emotions and how jealousy affects our interpretation of situations (*The Rationality of Emotion,* pp. 195–96).

8. For a discussion of spider and other phobias, see Calhoun.

but this, the objection continues, is implausible. To reply to this objection we need to consider the difference between trusting someone, however briefly, and having a relationship with that person that is predominately characterized by trust. In many circumstances, these will amount to the same thing, as, for example, when I trust a stranger in a momentary meeting: there is no relationship beyond the momentary that could be distrusting. Let us look at an example to see how this reply evades the objection.

Suppose that I have a friend who is particularly charming and particularly irresponsible. Time and time again she lets me down, and time and time again I forgive her and resume a relationship, promising myself that this time I will be more cautious, this time I will not count on her, this time I will remember to think of the ways in which I make myself vulnerable to her, and this time I will take measures to protect myself. I won't trust her again. For all my resolution, I might nonetheless find myself trusting her. It's true that whenever I become aware of doing so, I will resist the impulse and will once again be on my guard. At one extreme, I might only become aware of my having again trusted when I am again let down. I would say of myself that I find myself trusting her, even though, when I think about it, I'm aware that I shouldn't. Our relationship, for that time period, would have been characterized predominantly by trust. At the opposite extreme, my caution might undermine my tendency to view her with trust so that no sooner do I find myself viewing her that way than I call myself to attention and remind myself of all the reasons not to trust her: in this circumstance, I would not be willing to depend on her when the need arises. Here, I'm inclined to say that I don't trust her, although I fight the tendency to do so: I do not go far enough along with the patterns of salience and tendencies of interpretation that partly constitute trust for it to be the case that I trust her. Even when caution dominates, there is, though, the possibility of momentary trust—trust that is unnoticed and would be withdrawn as soon as it were noticed—and there will be momentary trust whenever I'm not quick enough to catch myself and reject the view of her that I have adopted. When I'm not quick enough at catching myself, my view of the other can give rise to the risk-taking behavior characteristic of trust. In this kind of case, our relationship is not one characterized by trust, but for all that, it can have moments of trust. Usually, when we say that A trusts B (within a certain domain), we mean that A's relationship with B (within the domain in question) is predominantly characterized by trust's distinctive way of seeing someone, and not merely that on occasion A sees B through the lens of trust. Thus, there is some truth in the claim that when I reject the appearances trust gives rise to, I don't trust, but this does not force us to say that trust requires more than a distinctive way of seeing someone; although a trusting relation-

ship (which is usually what we have in mind when we say that A trusts B) requires a consistent pattern of such interpretations. This solution to the objection is preferable to saying that trust requires us to have an endorsed attitude toward another, because we are generally not aware of our trusting and seldom bring it sufficiently clearly before our minds to endorse or reject it.

C. Clarifying the Distinction between Trust and Reliance

While trust is always a possible attitude to take toward a person, we sometimes rely on people instead of trusting them. So, for example, I can rely on someone to behave in a certain kind of way because I have evidence that it is likely that she will behave in that way out of, say, habit, fear, vanity, or stupidity. As Baier notes, trust is not a precondition for relying on someone.[9]

Trusting is not an attitude that we can adopt toward machinery. I can rely on my computer not to destroy important documents or on my old car to get me from A to B, but my old car is reliable rather than trustworthy. One can only trust things that have wills, since only things with wills can have goodwills—although having a will is to be given a generous interpretation so as to include, for example, firms and government bodies. Machinery can be relied on, but only agents, natural or artificial, can be trusted.[10]

Some cases of reliance are not grounded in perceived features of a person's psychology at all. Sometimes we adopt a policy of not checking up on people because to do so would be too time-consuming or too expensive. It's better to allow a few people to cheat on the coffee sign-up sheet than to devise a cheat-free method for collecting the

9. Baier, "Trust and Antitrust," p. 234.

10. It is a consequence of my account, though, that when we say nonnatural agents trust, our usage is analogous to our usage in attributing trust to a natural agent; but insofar as it is metaphorical to attribute affective states to nonnatural agents, the meaning is not precisely the same. (This is also true, though, when we attribute beliefs to nonnatural agents.) Sometimes government policies can enact something similar to the selective vision characteristic of trust, and the rationale for those policies can duplicate the expectation constitutive of trust. For example, a social welfare agency might decide not to use surveillance methods to eliminate cheating, on the grounds that the number of cheaters is likely to be small and can be further reduced by a policy of not checking since that would make the recipients feel they were being treated respectfully, and they would respond positively to such treatment. Here, the government agency would have been expecting its clients to respond favorably to the fact that they were being counted on. The agency's policy would have mimicked the way optimism gives rise to selective interpretation in that the agency would have proceeded on the basis of the assumption that cheating was not something to be expected. Note that the rationale for the policy matters: it would not be correct to say that the agency trusts its clients if they simply thought checking up on them was cost-inefficient.

coffee money. Devising a cheat-free method would simply take more time and cost more money than it would be worth.

There are some additional things that need to be said about the difference between trust and reliance. Sometimes we rely on things because we have no choice but to do so; thus we can be forced to rely on something when we are unable to predict that the event on which we rely will occur. However, if we have a choice about the matter, we will rely on someone only to the extent that we would be willing to make a prediction that the favored outcome will occur. In Section IV, we shall see that things are otherwise with trust: we can be justified in trusting even when we would not be justified in predicting a favorable action on the part of the one trusted. Our evidence for trusting need not be as great as the evidence required for a corresponding justified prediction. In this respect trusting is more like hoping than like predicting.

III. ADVANTAGES OF THE ACCOUNT

As we have seen, trust is to be distinguished from reliance in that trusting requires an attitude of optimism about the goodwill and competence of another as it extends to the domain of our interaction with them, and, in addition, trusting requires an expectation that the other will be directly moved by the thought that we are counting on them. It still needs to be shown that this account is adequate as an account of trust. An adequate account of trust should be able to explain at least the following three fairly obvious facts about trust: that trust and distrust are contraries but not contradictories,[11] that trust cannot be willed,[12] and that trust can give rise to beliefs that are abnormally resistant to evidence.[13] Because my account places an affective attitude

11. This is noted by Trudy Govier in "Trust, Distrust, and Feminist Theory," *Hypatia* 7 (1992): 16–33, 18.

12. This is noted by Baier in "Trust and Antitrust," p. 244.

13. This point is made by Judith Baker in "Trust and Rationality," *Pacific Philosophical Quarterly* 68 (1987): 1–13. Baker is concerned about the problem of reconciling trust with evidentialism, or the view that we should never believe anything without sufficient evidence. She claims that trust is "a kind of commitment, a state of the will" (p. 10). (If this were right, though, it seems trust should be able to be willed.) Trust still essentially involves beliefs, although it is to be assessed primarily for strategic rationality. She attempts to resolve the tension in saying that trust involves beliefs but is primarily assessed in terms of goal-directed, rather than truth-directed, rationality by pointing out the importance of trust for friendship: "But if a result of becoming someone's friend, of one's trust, is that barriers to honesty are removed and the other person is open with us, then trust in their veracity will be merited and end-directed rationality will not be opposed to truth-directed rationality" (p. 12). If believing makes it so, then the belief is justified. But, of course, believing does not always make it so, and so strategic and representational rationality won't always be in alignment. My account, which places an affective element at the center of trust, is able to finesse the evidentialist objection. I return to the issue of evidentialism in Sec. IV.

at the center or our understanding of trust, it is able to explain all these things.

Given that trust and distrust both involve attitudes, it should be the case that together they do not exhaust the possible stances we can take toward another's goodwill and competence. Optimism and pessimism are contraries but not contradictories; between them lies a neutral space. As a consequence, the absence of trust is not to be equated with distrust, for one may fail to trust without actively distrusting—one may simply not adopt any attitude at all toward the goodwill and competence of another. In between trust and distrust are found various forms of relying on and taking for granted which are not grounded in either optimism or pessimism about the other's goodwill.

Affective attitudes look toward features of the world that would make them justified and can no more be sincerely adopted in the face of a known and acknowledged absence of such grounds than a belief can be adopted in the face of a known and acknowledged lack of evidence. Because trust involves an affective attitude, it is not something that one can adopt at will: while one can trust wisely or foolishly, trust cannot be demanded in the absence of grounds for supposing that the person in question has goodwill and competence and will be likely to take into account the fact that one is counting on them. This is not to say that there can never be an element of decision in adopting beliefs or attitudes. We can, for example, *decide* that the evidence we now have is enough to support the belief, but we can't just decide to believe regardless of the evidence.[14] While trust cannot be willed, it can be cultivated. We cultivate trust by a selective focus of attention toward the grounds for trust and away from the grounds for distrust.

Trust gives rise to beliefs that are highly resistant to evidence. While affective attitudes can't be willfully adopted in the teeth of evidence, once adopted they serve as a filter for how future evidence will be interpreted. If I trust you, I will, for example, believe that you are innocent of the hideous crime with which you are charged, and I will suppose that the apparently mounting evidence of your guilt can be explained in some way compatible with your innocence. Of course this resistance to evidence is not limitless: given enough evidence, my trust can be shaken and I can come to believe that you are guilty. When my trust is shaken, I will come to see you in quite a different light: that certain shortness of temper that never seemed so important before, seemed always to be able to be explained away, now seems

14. Of course, that we cannot adopt an attitude in the face of a known and acknowledged lack of grounds is not to say that our affective stances follow our beliefs in a timely fashion, nor that we don't sometimes find ourselves experiencing "spill-over" feelings, as happens, for example, with phobias.

highlighted. I can come to see that, yes, you could have done what you are charged with, and, perhaps, even more strongly, that, yes, it is the sort of thing you would do. But in coming to see you in this way, without trust, I undergo a significant shift in the patterns of my attention with respect to your character and in my habits of interpretation of you, your character, and your motives.

If, as I have claimed, trust has an affective component and emotions are partly constituted by patterns of salience and tendencies of interpretation, it should come as no surprise that trust gives rise to beliefs that are highly resistant to evidence. For the same reason, trust and distrust have a tendency to seek out evidence for themselves and so to be, to a degree, self-confirming.

Bearing in mind these three facts that an account of trust ought to be able to explain, I turn now to Baier's account of trust in "Trust and Antitrust." According to Baier, trusting is a matter of *entrusting*. Trusting is analyzed as a three-place predicate: A trusts B with valued thing C.[15] Baier acknowledges that there are three difficulties with her account: It involves a degree of regimentation in that it may sometimes be difficult to specify exactly what is entrusted (p. 236). It might suggest a greater degree of consciousness and explicitness than our trusting relations typically display, so we need to guard against interpreting the model in this way (p. 240). And finally, it seems to overlook plain, non-goods-relativized, trust. But we might think that we should first trust before we entrust. (Baier herself notes this [p. 259] but thinks that we flatly rule out entrusting anything whatsoever to someone or some group of persons only when our interests are in complete opposition.)

I want to test an entrusting model against the three commonplace facts about trust mentioned earlier. It turns out that, because the model leads us to focus on the disposition of cared about objects at the expense of focusing on attitudes, it has problems explaining at least two of the three commonplaces; furthermore, with our attention drawn outward toward these objects, it is easy to lose sight of the crucial element of optimism about the goodwill of another. Trust becomes insufficiently distinguished from (mere) reliance.

It seems that one either entrusts valued thing C to B or one does not. If not entrusting is distrusting, then trusting and distrusting are contradictories. But it seems that while one has to either entrust or not, trust and distrust do not exhaust the options, and so trust and distrust are not to be equated with entrusting and refraining from entrusting. Explaining why trust and distrust should be contraries but

15. Baier, "Trust and Antitrust," p. 236. For the remainder of this section, page references in the text are to this article.

not contradictories is thus at least a prima facie problem for an account that analyzes trust in terms of entrusting. There may, though, be a way to preserve this distinction within an entrusting model. Perhaps trust involves a positive handing over of the thing entrusted, and distrust involves a positive *refusal* to hand over, a deliberate *withholding* of, the good in question, or, perhaps, purposive action to *protect* the valued thing. Trust and distrust could then be seen as contraries, for there is room for a neutral position in which one neither hands over the valued good nor holds it close to oneself. How good a reply this is depends on what it is that is being entrusted to another. There are three stances to take toward, say, you and the family silver: I may lend it to you, lock it up when I know you are visiting, or take no special precautions over it. It is less clear that there are three stances to take toward one's own self when walking down the street, and so it's not clear that this reply is fully satisfying.

The second commonplace that an account of trust must be able to explain is that trust cannot be willed. Baier notes this fact: "'Trust me!' is for most of us an invitation which we cannot accept at will—either we do already trust the one who says it, in which case it serves at best as reassurance, or it is properly responded to with, 'Why should I and how can I, until I have cause to?'" (p. 244). But *why* cannot one trust at will? If trust is entrusting it seems that I should be able to entrust at will, simply by handing over the relevant good. I may not feel very comfortable about it, but unless a feeling is built into the analysis of trust, that seems beside the point. Entrusting is an action and actions are, paradigmatically, things that can be willed. If, however, trusting involves an attitude, and attitudes cannot be adopted at will, we have an explanation for why one cannot trust at will.

Put just like this, the objection is surely unfair. We must cash out the "trust" as it occurs in the entrusting model's "A trusts B with valued object C." Perhaps without a belief in the reliability of the goodwill of another we cannot trust but instead can only rely on them. Baier says that the difference between trusting and relying on is that when we trust we rely on the goodwill of others toward us, whereas we may rely on others' "dependable habits, or only on their dependably exhibited fear, anger, or other motives compatible with ill will toward one, or on motives not directed on one at all" (p. 234). So we are to view trust as entrusting on the basis of a belief about the goodwill of the other. Trust is the action in conjunction with the belief that specifies its reason. Such a belief would have to be based on evidence and so could not be summoned at will. However, when an entrusting model is adopted, the significance of confidence in the goodwill of the other easily falls from view. This is because we can entrust where we don't, on my account, trust. That goodwill drops out of the picture when we focus on entrusting is shown in Baier's discussion of the moral

rightness or wrongness of trust relations. The cases she considers as trust relationships appear to lack this element of reliance on the goodwill of another: "Where the truster relies on his threat advantage to keep the trust relation going, or where the trusted relies on concealment, something is morally rotten in the trust relationship" (p. 255). I would suggest that in a situation such as this you haven't a morally rotten trust relationship, you haven't a trust relationship at all; instead you have a case of mere reliance. Optimism about goodwill is central to trust, but in situations like the one Baier describes, there is no goodwill. Trust does not seem to be sufficiently distinguished from relying on. Nor should this surprise us given that one can entrust where there is mere reliance. If I can rely on another's fear, my ability to control the purse strings, or the foolishness of another, I might be fully justified in entrusting them with something I care about, for I can know that they will not dare harm it or that it won't occur to them to do so. In such cases, confidence in some other aspect of a person's psychology has replaced confidence in her goodwill, but where this other thing is sufficient to ensure adequate performance, I need not also depend on her goodwill. This is not to say that relationships of reliance aren't sometimes mixed or can't depend on both kinds of elements.

Goodwill readily drops out of an entrusting model even when we attempt to include it, as Baier does. This is because we are led to focus on the disposition of cared-about objects rather than on attitudes toward a person, whom we might, as a consequence of holding such an attitude, willingly let get dangerously near things we care about. One of the chief motives for adopting an entrusting model—namely, that we be able to say what is in common and what is different between the various forms of trust, ranging as they do from trust in strangers to trust in intimates—can be accommodated within a nonentrusting model provided that we allow for variation in the domain over which the attitude extends. This lets us keep "plain trust" (p. 259), as an attitude directed toward a person and as explanatory of the kinds of risks we might willingly expose ourselves to with respect to that person, while yet being able to make the same kinds of distinctions that an entrusting model can.

An entrusting model is silent about the third commonplace a theory of trust should explain. And this is so even when such a model is fully spelt out so that trust is entrusting on the basis of a belief that the other has goodwill. The belief that another has goodwill may lead us, in the first instance, to be doubtful about her guilt. This is perfectly reasonable insofar as the evidence that supports the belief that the other has goodwill is also evidence for the belief that the person couldn't have done such a thing—think of character witnesses in criminal trials. However, as a belief of a perfectly ordinary sort, it should not be abnormally resistant to evidence, and it should not lead us to

hold additional beliefs that are themselves abnormally resistant to evidence.[16] But the beliefs we form on the basis of trust *are* abnormally resistant to evidence and so, in general, is the optimism about the goodwill of another that grounds such beliefs.

I conclude that an entrusting model does not sufficiently bring out the affective component in trust. In particular, it obscures the importance of optimism about the goodwill of another.

IV. JUSTIFIED TRUSTING

Given the usefulness of trust, should we say that a trusting attitude is the rational default position and that we should tend to approach the world with a trusting cognitive set? Or should we say that in the light of the harms to which we are vulnerable when we trust unwisely, the rational default position is one of distrust? Or, finally, should we say that the rational default position is one of neutrality?[17] Appropriate default stance is too sensitive to climate, and to domain and consequences as they interact to affect the expected disutility of misplaced trust, for there to be useful generalizations here. Further, for the individual truster, the appropriate default stance is linked to her assessment of tendencies in her own trusting and distrusting.

In climates in which there is strong motive to be untrustworthy, it would require more evidence for our trust to be justified than in climates where there is little incentive to untrustworthiness. A final verdict on whether a particular act of trusting is justified will have to step beyond that particular case to examine general features of the social climate we inhabit. Thus I take it that at the height of the Chinese Cultural Revolution, the justified default position was one of distrust, and that it took more evidence to be warranted in moving from this position than it would take to warrant justified trusting in a more favorable climate. During those campaigns, people accused of being counterrevolutionary were subject to public shaming, beating, and incarceration. At the time, there was strong motive for people to be untrustworthy. Informers were held up as model citizens, and anyone who displayed goodwill toward someone who became a target of the campaigns was in danger of herself becoming a target. In such a

16. In Baier's newer work on trust, the entrusting model is less emphasized, and she acknowledges that trust also involves an affective aspect, though she does not attempt a detailed account of it (*Tanner Lectures on Human Values*, pp. 111–12). For Baier, the affective aspect is not central in an account of when trust is justified. In Sec. IV, I will claim that the affective aspect of trust is central for understanding when trust is justified.

17. Talk of default positions is compatible with the Sec. III claim that emotions look to the world for evidence. Consider anger: we can admit that anger can't be willed but still inquire whether it is better to be irascible, placid-tempered, or something in between.

climate, showing goodwill toward others was dangerous. In contrast, in an ideally moral climate, the interests of each would be harmonized so that trust could flourish. The motivation to be untrustworthy would diminish, and there would thus be grounds for the expectation that those we encounter are trustworthy.

Domain and consequences interact to determine which default stance is justified and how much evidence we need to move from that default stance. If I am to have the depth of trust that would make it reasonable to entrust you with a secret of mine, then I'll want to have quite a bit of evidence about your character. If, though, I am to trust you not to attack me in the street, I may need no particular evidence about your character at all. This might seem to be a counterintuitive result, since surely it is worse to be attacked in the street than it is to be embarrassed by a confidence indiscreetly betrayed. However, domain is here signaling likelihood of performance. We are all aware of the lively pleasures of gossip and of the strength of character required to resist them. In contrast, it is not hard to refrain from harming a stranger on the street; that just takes basic decency, a trait that we can assume is widely shared, unless the climate is sufficiently bad. Once we hold domain fixed, consequences become of the first importance: of course I'm going to need more evidence of your trustworthiness before I willingly tell you a secret that, if spread abroad, would be damaging to me, than before I tell you a secret whose disclosure would be merely embarrassing. It is not, therefore, that one or the other of domain or consequence is always the most important; rather what is important is how they interact to determine the expected disutility of misplaced trust (or distrust).

While climate, domain, and consequences are variables determining which default stance is justified that extend across agents, the fourth variable determining the appropriate default stance is agent-specific. Some agents have reason to be distrustful of their tendencies toward trust in certain domains. When we believe that we are poor affective instruments, either in general or across a specific range of cases, we should distrust our trust, or distrust our distrust, and demand a correspondingly higher amount of evidence before we let ourselves trust or distrust in the kinds of cases in question. Consider responses to physicians. We can imagine someone with a tendency to find authoritative and avuncular physicians trustworthy and physicians who acknowledge the tentativeness of their diagnoses and the limits of their art untrustworthy. Given how sexism shapes what we take to be signs of competence, we should be wary of our tendency to trust when an etiology of that trust tells us it is as likely to be caused by mannerisms of privilege as by marks of trustworthiness.

Because climate, domain, consequences of misplaced trust, and appropriate assessment of the tendencies of our own trusting and

distrusting affect how much evidence is needed before our trust can count as justified, the question of the rational default position has no general answer. However, there is still an important question to be addressed: Are there any instances of apparently justified trust or distrust where we would not want to say that the person would be justified in having the belief that the other was trustworthy or untrustworthy? If there can be such cases, then if we advocated an analysis of trust which made trust fully or partly constituted by a belief about the other's trustworthiness, we would be forced either to reject evidentialism or admit that the cases weren't justified after all.

There are two places to look for examples of trust leaping ahead of the evidence: when trust is governed by forward-looking or instrumental considerations, and when trust is governed by backward-looking considerations of evidence but our responses seem to outstrip the evidence. Let's examine the forward-looking cases first.

Earlier I remarked that trust cannot be willed but that it can be cultivated. We might want to cultivate trust toward people in general, toward members of a certain group, or toward a particular person. Moreover, it seems that we can sometimes be justified both in attempting to cultivate trust and in the trusting that is the result of such cultivation. If trust and distrust are partly constituted by patterns of attention, lines of inquiry, and tendencies of interpretation, it should be possible to cultivate them by controlling our patterns of attention, our lines of inquiry, and our interpretations. Thus, while trust cannot be directly willed, we can will to pay attention to the kinds of things that are likely to support, create, or extend our trust, and we can will to refrain from focusing on the kinds of things that are likely to undermine and limit our trust.

Sometimes we set about cultivating trust because we think that by trusting, and displaying our trust, we will be able to elicit trustworthy behavior from the other. When we do this our hope is that by trusting we will be able to bring about the very conditions that would justify our trust. It might be thought that we do not need to inquire whether attempts at this sort of bootstrapping can be justified, for we need never actually trust on the basis of forward-looking considerations—all we need do is act *as if* we trusted. To actually set about trying to trust is to do more than is needed. It is a mistake, though, to think that acting as if you trusted will have the same results as acting on the basis of genuine trust, cultivated in the hope of bringing about trustworthiness. Acting as if you trusted and genuinely trusting could have the same result only on the assumption that there is no perceptible difference between the behavior that would be produced from trust and the behavior that would be produced from acting as if you trusted. But this assumption is implausible in the kinds of cases where one is most likely to adopt this sort of strategy in the first place. Trusting in

the hope of eliciting trustworthiness is a pointless strategy to adopt with those with whom we have infrequent contact. In such circumstances our strategic trusting could not bear fruit. Instead, it is the kind of strategy a parent might use with a child, or a lover with her beloved. But it is precisely the frequency of contact between the one who would trust and the one she would elicit trustworthiness from that makes it implausible to suppose that merely acting as if you trusted could, on each of many separate occasions, result in behavior indistinguishable from the behavior of one who genuinely trusts. If this is so, then we do need to ask when, if ever, bootstrapping is justified.

Bootstrapping is not always possible and not always reasonable. It won't be possible if we cannot find sufficient foundation in evidence for our trusting. Despite our attempts to control our patterns of attention and our interpretations, we might be unable to find enough to focus on to support our trust. Our attempts at giving positive reinterpretations of those aspects of a person that might otherwise have tended to support the hypothesis that she is untrustworthy have the feel of fantasy and wish fulfillment. They do not ring true. Whenever trust can be achieved only through a fantasy construction, our trusting is unlikely to elicit trustworthiness from the other, for if fantasy is required to see the other as trustworthy, it is highly unlikely that the other has the potential for trustworthiness.

In addition to cultivating trust in order to elicit trustworthiness from another, we sometimes cultivate trust in order to realize a conception of ourselves. So, for example, the rape victim whose trust in others has been shattered might set about cultivating trust because she sees herself as someone who is free-spirited and bold, and she does not wish to be the kind of person who is timid, protective of the self, and on the lookout for betrayal. She does not want her horrible experience to lead to a change in herself. The trust that results from willful cultivation can be rational. Provided that its cultivation did not require fantasy and distortion, it can be reasonable to view it as keyed to real and perceptible features of the agent's situation.

The second sort of cases involve flash intuitive assessments that do not seem to be based on evidence sufficient to support a belief. Let us consider an example taken from Greenspan's *Emotions and Reasons.*[18] I find myself feeling suspicious of a salesman, worried that he will harm my interests, worried that he is not trustworthy. Let us suppose,

18. Patricia S. Greenspan, *Emotions and Reasons* (New York: Routledge, 1988). Greenspan argues that suspicion can be justified when one would not be justified in forming a belief that the other is untrustworthy. Her account of justification also stresses forward-looking conditions. However, she does not think that skepticism about our own capacity as an emotional instrument is especially undermining of the justifiedness of an emotion.

further, that the salesman has been recommended to me by a friend whose judgment in such matters I believe to be reliable. On the basis of this recommendation, I believe that the salesman is trustworthy, yet I find myself unable to help viewing him with suspicion. I continue to see him as untrustworthy, although I am not yet prepared to abandon my belief that he is trustworthy. I cannot articulate why I view him with suspicion, except to say that there is something creepy about him, something in his manner that I don't like. Finally, let us suppose, although I don't know this myself, that I am not, in general, a reliable detector of untrustworthiness. My suspicion would not track untrustworthiness across a suitable range of counterfactual circumstances relevantly similar to the present one. Thus, as the example is set up, if having a justified belief requires being able to give an account of what justifies that belief, I haven't got a justified belief. Similarly, if having a justified belief requires having a belief that tracks the truth across some range of counterfactual circumstances, I haven't got a justified belief. And if having a justified belief requires having a belief formed by a reliable process and the absence of undermining beliefs, I haven't got a justified belief, for my belief that the salesman is trustworthy appears to undermine my perception of him as untrustworthy.[19] Thus, it seems that I wouldn't be justified in forming a belief that he is untrustworthy on the basis of my seeing him as untrustworthy.

If we think that trust and distrust are primarily beliefs, it seems that—regardless of any of the variables mentioned earlier—we would have to say that my distrust could not be justified. But it seems to me that, especially if the stakes are high, I might still be justified in following through with the lines of inquiry and patterns of salience that are constitutive of distrust.[20] This is because emotions and other affective states often do represent the world in the way it is: those we are suspicious of often are untrustworthy.[21] I do not mean to claim that distrust would have to be justified in cases of this sort, only that even though we've *decisively* shown the belief that the other is untrustworthy is unjustified, we haven't decisively shown that distrust is unjustified. To do that we would have to step back and examine the other variables affecting the justification of trust, and they could well return the verdict "justified." This example better lets us understand the importance

19. These three options are meant to exhaust the possible accounts of what makes a belief justified. The first is an internalist account and the others externalist accounts.

20. That is why this sort of example is best developed with cases of suspicion, since usually the costs of being wrong in our distrust are less than the costs of being wrong in our trust. But this is a generalization that admits of exceptions: there can sometimes be severe consequences of misplaced distrust.

21. In *Passions within Reason* (New York: Norton, 1988), Robert Frank presents an evolutionary argument for why emotions can have this role.

of the truster's assessment of the tendencies in her own trust and distrust: the metajustification in terms of the worth of following up on affective appearances would not be available to those with reason to distrust their distrust.

If there can be cases of the sort I have described, then an account of trust that makes affect central has an unexpected payoff: it lets us say that such cases can be justified without confronting the evidentialist thesis. It is beyond the scope of this article to argue the merits of evidentialism. It might well be that evidentialism is false—perhaps, for example, we can be justified in believing on instrumentalist grounds. But if, as I have argued, trust and distrust are not primarily beliefs, then trust and distrust cannot be used to unseat evidentialism.[22] Equally, though, evidentialism cannot be used to challenge our intuitions about when trust and distrust are justified.

22. In "The Virtue of Faith," *Faith and Philosophy* 1 (1984): 3–15, Robert Adams claims that trust requires beliefs that go beyond the evidence.