

Deciding to Trust, Coming to Believe¹

RICHARD HOLTON, UNIVERSITY OF EDINBURGH

If you have ever taken a drama course, you have probably played this game. You are blindfolded. You stand in the middle of a circle formed by the others. They turn you round till you lose your bearings. And then, with your arms by your sides and your legs straight, you let yourself fall. You let yourself fall because the others will catch you. Or at least that is what they told you they would do. You do not know that they will. You let yourself fall because you trust them to catch you.

If you are like me, there is a moment at which you weigh up whether or not to let yourself fall. How does it feel at that moment? It feels as though you are *deciding* whether or not to trust. I think that we should take this feeling at face value: there are circumstances in which we can decide to trust.

Does my decision to trust the others entail that I *believe* that they will catch me? If it does, does this in turn mean that when I decide to trust them, I also decide to believe that they will catch me? I think not. In order to trust I do not need to believe. Certainly my trust involves a certain state of mind; but it does not seem to me that it always involves a belief. Mightn't I be most uncertain that I will be caught, but decide to trust anyway? Mightn't I decide to trust thinking it as likely as not that you will let me fall?

Let me give another example of deciding to trust, one that comes more clearly from the realm of the moral. 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; perhaps you simply think it is the way you ought to treat one of your employees.² Of course your belief about the likelihood that he will steal will be one factor in your decision whether to trust. It might be that if you really believe he will steal, you will not be able to trust him. But you can trust him without believing that he will not. these states of belief are not clean negations of one another

Trust, I will suggest, is a distinctive kind of attitude involving a distinctive state of mind. My project is to look at the ways in which it is distinctive, and the ways in which it interacts with belief and with the will. Since my focus is on trust as a state of mind, there are a host of questions that I will not properly address: questions about when, and to whom, trust is appropriate. This is not, of course, because I think them unimportant; on the contrary, I think them too important for the sketchy treatment I could afford them here.

We see attitudes similar to trust in other places. H. G. Wells wrote that riding a bicycle is like conducting a love affair: it is, he said, a matter of faith.³ He went on to speak as though faith is a matter of belief. Believe and you can do it, doubt and you cannot. But I think that this rather mis-states the insight. To ride a bicycle you do not

¹ Thanks to John Campbell, Bas van Fraassen, Lloyd Humberstone, Rae Langton and two referees for comments on drafts of this paper.

² The idea that trusting someone can be a way of making them trustworthy, that, as he puts it, trust "can generate the very behaviour which might logically seem to be its precondition" has been noted by Diego Gambetta in 'Can We Trust Trust,' in Gambetta (ed.), *Trust: Making and Breaking Cooperative Relations* (Oxford: Basil Blackwell, 1988) pp. 213-37, at p. 234.

³ H. G. Wells, *The Wheels of Chance* in *The Atlantic Edition of the Works of H. G. Wells* (London: T Fisher Unwin Ltd, 1925), Vol. 7, p. 19.

main thesis
trusting \rightarrow believing

not about
the morality or
rationality of trust

a sort of rehabilitative
trust

need to believe that you can do so. You need to act as you would act if you did believe. You need to get on the bike and push off confidently. All you need is to be in a state of mind that leads you to do this. Perhaps for some people the only state that will do the job is the belief that they will succeed; but this is not universally so. The faith of which Wells speaks is not, I think, belief.

A more familiar example of much the same idea comes from religion: if you have faith, some say, all else (including belief in what you took to be incredible) will follow. Here we should not let the bicycling example lead us astray. The Christian writer who says that all you need is faith need not be proposing the religious equivalent of simply pushing off on the bicycle. That is, she need not be proposing a behaviouristic solution: act outwardly as a Christian would, and in time the right state of mind will come. In the case of the bicycle there is nothing more to faith than having any state of mind that leads you to push off. Whereas acquiring religious faith involves more than being in a state that leads you to act like a Christian. It involves a particular state of mind: something much closer to belief. Nevertheless, the conception of faith that is involved here is not quite the conception of a belief. In the first place, it is too immediately under the control of the will; in the second, Christians speak too frequently of the idea that belief will *follow* from faith.

This is no paper in theology, so I shall not pursue the nature of Christian faith any further. But the distinction between the simple state of mind that is involved in having faith to ride a bicycle, and the much more complex state that is involved in having faith in God, finds an echo in the cases of trust I will discuss. We will need to distinguish the relatively simple attitude that is involved when I decide to trust you to catch me, from the complex attitude when I decide to trust the things you say. I shall come to the complex attitude in the last section; let me start with the simple one.

analogy for discussions on faith

Trust and Reliance

What is trust? Annette Baier has given an influential account.⁴ According to her, trust must be distinguished from reliance. It is a special kind of reliance: reliance on a person's goodwill towards one. Why the distinction? Because, Baier claims, there are times when we rely but do not trust:

We may rely on our fellows' fear of the newly appointed security guards in shops to deter them from injecting poison into the food on the shelves, once we have ceased to trust them.⁵

I agree that we need to distinguish trust from reliance; but I do not think that Baier's way of doing so is quite right. In the first place, relying on a person's goodwill towards one is not a sufficient condition for trust; the confidence trickster might rely on your goodwill without trusting you. Secondly, it is not a necessary condition: I can trust a person without relying on their goodwill towards me. I can, for instance, trust someone to look after a third party or to look after themselves, without requiring that they have goodwill towards *me*. One member of an estranged couple between whom there is precious little goodwill can still trust the other to look after the children. Perhaps then the point is that the person I trust must have goodwill towards the *object* of the trust:

the "confidence trickster" objection to Baier's goodwill-based account of trust

⁴Annette Baier, 'Trust and Antitrust', *Ethics* 96 (1986) pp. 231-60. Since writing this paper I have seen a copy of Baier's Tanner lectures, 'Trust', *Tanner Lectures on Human Values* 13 (1992) pp. 107-174. These do not seem to modify her original account substantially; they do however include a discussion of a case very much like my drama class example (p. 115).

⁵ 'Trust and Antitrust' p. 234.

towards the thing they are trusted with. That is not right either: I can trust someone with my car whilst knowing they have no goodwill towards *it*. So perhaps we could disjoin the two accounts: say that when I trust I rely on there being some good will somewhere, either towards me, or towards the object of the trust. But this still is not right. Consider one of Baier's own examples: we trust our enemies not to fire at us when we lay down our arms and put out a white flag.⁶ There is not necessarily much goodwill there. Our enemies' restraint might be grudging, driven simply by a belief that it is wrong to shoot someone who has surrendered. To impute goodwill here would be to deprive the notion of all content. Nevertheless, talk of trust is not out of place.

where to 'place' the goodwill in Baier's account? It gets confusing

We need to take a rather different approach if we are to distinguish trust from reliance. In fact, I think we need to make a number of distinctions. As a preliminary, let me distinguish two different states of reliance: relying on something happening; and relying on a person to do something. When I rely on something happening I need not believe that it will happen. But I do need to plan on it happening: I need to work around the supposition that it will.⁷ Relying on a person to do something requires more. Consider again Baier's example of the would-be poisoners. Baier says that it involves reliance, and so it does: we rely on the would-be poisoners not getting to the food. But would we say that we rely on them not to poison us? I think not. To see why this is so, consider a related case. Suppose that the would-be poisoners are caught, convicted, and imprisoned. We can rely on them staying in prison for their term; but it is not true that we rely on them to stay in prison. To rely on a person doing something is not just to rely on a certain state of affairs happening: the state of affairs in which they do that thing. Rather, it is to rely on them doing it from a motivation that stems in some way from them. If I rely on you to stay in prison, I do not simply plan on the supposition that you will stay there; I plan on the supposition that you will stay there because you are motivated to stay there, and not just because you have no choice.

reliance on people doing things involves more than reliance on something happening

The same point applies in the case of the would-be poisoners. Why do we not say that we rely on them not to poison us? The reason, I think, is that their motivation is insufficiently self-generated.⁸ Admittedly in this case they have some motivation to desist: fear of getting caught. Unlike the prisoners, they are not physically barred from starting on their plan. But this motivation is insufficiently internal in its origin. Exactly what does count as an internally driven motivation is not clear to me; fear that frustrates one's plan surely does not.

What, in the light of this, has become of the distinction that Baier drew between trust and reliance? She claimed it was necessary to make that distinction because in cases like the one she described we rely but we do not trust. Yet we have seen that the only distinction we need in order to describe the case is that between relying on something happening and relying on a person to do something.

Nevertheless, it is surely right that there is a difference between merely relying on a person to do something, and trusting them to do it. What is special to those cases in which we trust? Baier herself raises what I take to be the central point, but I think she rather misconstrues it. In cases where we trust and are let down, we do not just feel disappointed, as we would if a machine let us down. We feel betrayed. Baier takes this as evidence that when we trust someone, we rely on them to have goodwill towards us.

another distinction

⁶ Ibid.

⁷ Is this the same as acting as if I believed it would happen? In one sense yes: in certain circumstances my behaviour will be as it would be if I believed. But belief and reliance involve different dispositions, as I discuss below.

⁸ I assume here that we should give the same account of relying on someone *not* to do something as we gave to that of relying on them to do it.

But as we have seen, that is not always the case. The central point is rather that betrayal is one of those attitudes that Strawson calls the *reactive attitudes*.⁹

These are attitudes that we normally take only towards people. We feel hurt or resentful when they let us down; grateful, perhaps touched, when they help. In contrast, when a machine breaks down we might feel angry or annoyed; but not (unless we are inveterately anthropomorphic) resentful.

Resentment and gratitude are examples of the particular attitudes that we feel towards people when they act in certain ways. Behind them stands a more general attitude, which Strawson calls the *participant attitude*, but which I shall call the *participant stance*. Obviously not everyone we interact with does something to provoke a reactive attitude, whether positive or negative. Many of our dealings are neutral in this respect. Nevertheless, if the people in these dealing were to hurt us, we would typically feel resentment; and if they were to help, we would typically feel gratitude. We are ready to take particular reactive attitudes should they act in certain ways. In having such readiness we take the participant stance towards the people concerned. The readiness is partially constitutive of the stance (partially, since taking the stance can require engaging in a whole network of further attitudes and actions, and perhaps beliefs). We might not actually take the appropriate reactive attitude even if it is deserved; we might, for instance, forswear resentment by forgiving someone. But there is a difference between forswearing resentment, and having no readiness to feel it. The former might be a response to our friends; the latter an attitude to a machine.

I think that the difference between trust and reliance is that trust involves something like a participant stance towards the person you are trusting. When you trust someone to do something, you rely on them to do it, and you regard that reliance in a certain way: you have a readiness to feel betrayal should it be disappointed, and gratitude should it be upheld. In short, you take a stance of trust towards the person on whom you rely. It is the stance that make the difference between reliance and trust. When the car breaks down we might be angry; but when a friend lets us down we feel betrayed.

Perhaps it is best to see the stance of trust as part of the participant stance, in the sense that trusting someone is one way of treating them as a person. But if this is right, it shows how important it is that we do not treat the participant stance as an all or nothing affair. Even when you do trust a person, you need not trust them in every way. Trust is a three-place relation: one person trusts another to do certain things.¹⁰ You can trust a person to do some things without trusting them to do others.

⁹P. F. Strawson, 'Freedom and Resentment', in his *Freedom and Resentment* (London: Methuen, 1974) pp. 1-25. I can only give a brief sketch of Strawson's discussion here. Readers in search of more might consult, in addition to 'Freedom and Resentment', J. Bennett, 'Accountability', in Z. van Straaten (ed.), *Philosophical Subjects* (Oxford: Oxford University Press, 1980) pp. 14-47; Christine M. Korsgaard, 'Creating the Kingdom of Ends: Responsibility and Reciprocity in Personal Relations', in J. Tomberlin (ed.), *Philosophical Perspectives 6: Ethics* (Atascadero: Ridgeview, 1992). Note that Strawson initially characterizes the reactive attitudes as resulting from our concern that others have goodwill towards us (p. 5). So it might seem that by appealing to the reactive attitudes I am bringing in by the back door that very feature of Baier's account that I took issue with: the idea that in trusting someone we rely on their goodwill towards us. But by the following page Strawson has relaxed this, I think rightly, to the claim that the reactive attitudes result from our concern about the goodwill *or regard* of others.

¹⁰Baier also thinks that we should treat trust as a three place relation. But she suggests that we take as central the idea of one person trusting another with some valued thing ('Trust and Antitrust' p. 236). I think that this is a special case of the more general three place account that I propose: the special case in which the act that we trust the other to perform is that of looking after something.

Of course, this gives nothing like a reductive analysis of the stance of trust. It does, however, help us to see how trust fits in to a larger picture; and it provides an explanation of some points that we have already noted. Firstly, it explains why the confidence trickster who relies on our goodwill need not be said to trust us. He takes no participant stance towards us; he treats us simply as objects to be manipulated to his advantage. Secondly, it helps in elucidating the connection between trust and belief. Strawson contends that to take the participant stance towards someone is not to form a belief about them. It is to take a practical stance. The same goes, I think, for trust. Trusting someone does not involve relying on them and having some belief about them; a belief, perhaps, that they are trustworthy. What it involves is relying on them to do something, and investing that reliance with a certain attitude. This is to take a practical stance.

I need more on what this term means in philosophy in general

Strawson does not claim, however, that there is *no* connection between our beliefs about a person and our adoption of the participant stance towards them; he says that our beliefs about them are important in determining whether we think it appropriate to hold them responsible. In the case of trust at least, I want to go further. I do not simply think that our beliefs about a person will influence whether or not we adopt a stance of trust towards them; in addition, I think that adopting the stance will in turn lead us to form various beliefs. This is not because the stance logically entails the beliefs; a stance cannot enter into entailment relations in that way. Rather it is because unless we formed the beliefs, we would not really have taken the stance. I will develop this point further in the final section. First, let me say something to quieten the worry that I have left no space for the distinction that I want to draw between belief and states like trust and reliance.

The point will be clearer if I focus on the simple state of reliance. I have said that when one relies on something to happen one works the supposition that it will happen into one's plans; and I have claimed that such a state should be distinguished from belief. But it might be objected that I have left no room for such a distinction. According to the objection, your belief that a certain thing will happen just is the disposition that you acquire when you work the supposition that it will happen into your plans. One response to this is to say that whilst belief and reliance might both involve dispositions, they each involve some more: a distinct feel, which is not manifested in the disposition, but which is good enough to tell the two apart. Perhaps this is right; but I think the objection can be faced directly. Even if we accept the dispositionalist account of mental states, belief and reliance can be distinguished.

To see this, consider first what the dispositionalist would say about lying or other forms of pretence. Suppose I live in a society where those who do not embrace the one true faith are harshly dealt with. Knowing this, and knowing that I have no chance of changing the society, I pass my life acting as one who truly believes. When I decide on pretence I take on a complex disposition: the disposition to act as a believer would. But I do not believe. Those who see belief as a disposition must have some way of distinguishing belief from perfect pretence. What they can point to is the fact that belief and pretence differ in their causal genesis. One is caused by convincing evidence; the other by pragmatic necessity. And this in turn means that they are not the same disposition. In circumstances in which there is no need to feign belief, the pretender gives up pretending; but the true believer goes on as before.

*response to an objection:
belief and reliance involve
different dispositions and thus
have to be different*

I say much the same about reliance. As the fake believer has reasons for her pretence, so we have reasons for reliance. We can show how the disposition of reliance is different from the disposition of belief by examining the cases where those reasons no longer obtain. I can rely on a rope that I doubt to be secure because I have no

alternative; but the disposition that I exhibit in so doing is different to that which I would exhibit if I believed it were secure. If I no longer need to rely on the rope, if I were offered an alternative I believed to be secure, I would abandon my reliance on the rope and take the alternative. If I were confident in the rope I would not.

I have reason for simple reliance on an object if I need something done and reliance on it is my best bet for getting it done; likewise for simple reliance on a person. But in cases which involve not just simple reliance but trust, my reasons can be more complicated. Just because trust involves moving to a participant stance, I can have further reasons to trust, since that move can itself be something I value. Suppose we are rock climbing together. I have a choice between taking your hand, or taking the rope. I might think each equally reliable; but I can have a reason for taking your hand that I do not have for taking the rope. In taking your hand, I trust you; in so doing our relationship moves a little further forward. This can itself be something I value. We need not imagine that you would be hurt if I chose the rope over your hand; you might be perfectly understanding of the needs of the neophyte climber. But our relationship would not progress.

choosing trust as a way to
further relationships

Consider again the drama class example. When I let myself fall, and do not put out my hand to save myself, I am relying on you to catch me. It could be that that is all I am doing: I might be relying quite without trust. I might hate the drama class, and regret the day I enrolled for it. I might find you, my classmates, utterly incomprehensible. I might rely on you to catch me because this is the only way to complete the course and get the credit I need. But should you fail to do so I would feel no sense of betrayal; just a grim confirmation of your alien ways.

More likely, however, if I let myself fall I will trust you to catch me. This could be because I think that an atmosphere of trust is important for a drama class; and that one way to work towards such an atmosphere is to trust you. I could value the relationships that my trust will bring. But even if I do not, even if my aim is simply to get through the course, I could still have reason to trust you. You might have explicitly invited my trust; or perhaps I think that you have done so implicitly by participating in the game. And, given that I already have reason to rely on you, I can think that your invitation to trust gives me a reason to take an attitude of trust towards you. If you let me fall now I will indeed feel betrayed. Spelling my reasons out like this makes them seem calculated. We should not think, however, that in general this is so. My reasons might be hidden or unthinking. I might have always trusted; perhaps I can imagine no alternative but to go on.

Trust and the Will

As I have said, I think that we can decide to trust. More precisely, I think that in some circumstances we can decide to trust. I do not say that every case of trust is a case of a decision to trust. Nor do I say that we can decide to trust just anybody to do just anything. We are constrained in the trust that we can extend; but within the constraints there is some room for choice. Since my account of trust involves two parts, reliance and a stance of trust, you might think that there are two places where choice might come in: at the level of the reliance and at that of the stance. The idea of deciding to rely presents few problems. Just as the non-believer in the 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. But what of taking the stance of trust; do I have any control over that?

Sometimes, as Strawson points out, I have some control over my reactive attitudes. To take an example already mentioned: if I decide to forgive you, I decide to renounce

my feelings of resentment. Perhaps sometimes I can go further, and decide to suspend the participant stance towards you altogether. When things get too much, I can try to think of you as a hopeless victim of your upbringing, and sometimes I might succeed. But it is surely not right to think that I can decide to rely on you, and then make a further decision about whether to treat that reliance as trust. The mistake there is to treat reliance and trust as two quite different stages in my attitude towards you. They are not. Typically, to decide to rely on a person is to decide to trust them, since in relying one automatically takes the stance of trust.

There are, however, circumstances in which this is not so. The cases fall into two sorts (which parallel the two sorts of cases in which Strawson says we will not hold a person responsible for their actions¹¹). Firstly there are those cases in which I do not take the participant stance towards the person at all. Here we have a general suspension of the stance of trust. Secondly there are limited suspensions of the stance of trust, concerning particular instances of reliance. Here the general stance of trust towards the person remains in place. An obvious case of this can come when I rely on you without your knowing, or being reasonably expected to know, that I am doing so. Here I would not feel betrayed should you let me down. A more complicated case might arise if I rely on you to do something which I know to be well beyond what I could reasonably expect of you. Here again I might not feel aggrieved if you let me down. It is not that I am ready to forgive you; rather that although I relied on you I did so without the kind of emotional seriousness that is needed for trust. I see your failure as no reflection on you. Of course, if I take this attitude too frequently, if I set my expectations of your capacities too low, this case can shade off into one in which I do not fully take the participant stance towards you. And your reaction might be one of hurt rather than relief.

In "Trust and Antitrust" Baier insists that we can never decide to trust.¹² Why is this? It is not obvious that the claim is entailed by her own account. Indeed, she speaks at one point of trust as "accepted vulnerability to another's possible but not expected ill will (or lack of good will) toward one."¹³ Barring the talk of good will, that sounds rather close to the kind of account I have been proposing here: surely *acceptance* is under the control of the will. Perhaps the explanation is that she is concerned less with the kind of trust of which I have been speaking so far, and more with what I shall call a *trusting relationship*. I have treated trust as a three-place relation holding between two people and an action. Contrast this with a two-place relation like friendship or hatred: if I am really your friend, I am your friend *simpliciter*. I am not your friend *for* bus driving, or *for* clothes buying, or whatever. Nevertheless, it seems that there is often a two-place relation that stands behind cases of trust. I have a background relationship of trust with some people; and it is in virtue of this that I would trust them to do a whole range of things for which they have not yet proven their trustworthiness. Can I choose to have such a relationship? Surely not. Compare the case of friendship. I choose my friends; but I cannot choose to be a close friend of yours straight-off. I can choose to start along the path that will lead to that relationship; and I can control, to some extent, how quickly we move along it. However, close friendships are things which develop over time. Similarly with trusting relationships. If you and I trust each other in various ways over time, and our trust is not betrayed, we will be likely to build a trusting relationship. That is not to say that there will be some particular thing that I will trust

deciding the extent of one's participant stance

disagreeing w/ Baier, who rejects the idea that trust can be decided upon

sensical for the "two-place predicate" relationship kind of trust

¹¹'Freedom and Resentment' p. 7.

¹²Pp. 244-5. In her later Tanner Lectures she has changed this to the more moderate claim that "trusting is rarely something we *decide* to do" (p. 123).

¹³'Trust and Antitrust' p. 235.

you to do. Rather it is to say that I will in general be more ready to trust you: partly because I am confident that you will not betray that trust, and partly because, having trusted you before, further trust becomes appropriate. A trusting relationship makes a greater range of trust available to me. As with a close friendship, so with a trusting relationship: I cannot simply decide to have one. But that does not mean that I cannot sometimes decide to trust.

This leads to a further consideration. I have been treating trust as, so to speak, a one-sided relation. I have spoken of trust as an attitude which can be taken towards someone without their inviting it, an attitude which can be unreasonable, even pathological. Perhaps you will think that this is a mistake; perhaps you will think that the only true trust is that which is appropriate, made in the context of a suitable relationship. If so we have been thinking of rather different things. What you think of as trust, I think of as appropriate or well-founded trust. I shall stick with my usage, since I think it both more natural and more useful. But might a divergence here explain why I think trust can be controlled by the will, while others do not? I think not. Certainly the circumstances in which I can decide to trust *and* where my trust will be appropriate, are rarer than those in which I can simply decide to trust. Nevertheless, I think they exist: fill in the details in the right way, and the drama class and shop assistant examples should each be enough to show that.

key to Hollan's formulation of trust
is this idea that it can be taken up
irrationally or unreasonably
↓
this is not why he thinks we have
willful control of trust

So what are the constraints on trust: whom can I decide to trust, to do what, under which circumstances?¹⁴ Questions of this sort lead us into an extremely difficult area. There is a certain class of impossibilities that loom large in everyday life, but which have been little discussed in analytic philosophy. Shyness, fear, anger: each limits what is possible for us in distinctive ways. Moreover, these limits are very personal. What is possible for one person is not possible for another, even in the same circumstances. When we ask whom we can bring ourselves to trust we ask what is possible and what is impossible in this sense. And I do not have very much to say.

What I do have to say falls into two parts. The first concerns the attitude of reliance in general; in particular, the relation of reliance to belief. I have argued that we can rely on a person to do something without believing that they will do it. But what if we believe that they will not do it? Perhaps I think it slightly more likely than not that you will fail to catch me. Here I can still rely on you to catch me. But what if I am convinced that you will let me fall. Now it seems likely that I simply will not be able to rely on you.

So, there is a certain asymmetry concerning the requirements on my beliefs. I do not need to have the belief that you will do what I rely you to do, but I do need to lack the belief that you will fail. Is there any explanation for this? We can give one by invoking a consistency requirement on reliance. It is a reasonable requirement that my reliance be consistent with my belief: if I rely on you to perform some action it must be possible for you to perform it and for my other beliefs to be true. When I rely on someone to do something, I work this reliance into my plans: I plan on the supposition that they will do it.¹⁵ But if I believe that they won't do it, then that too is something I should work into my plans. So if I rely you to do something but believe that you won't,

consistency requirement for
reliance w/ belief

¹⁴Beware of one use of the phrase "Whom can I trust?" in which it means "Whom should I trust?" That's not what I mean by the question (although clearly your beliefs about who you should trust will restrict who you can trust in the sense which interests me).

¹⁵Contrast this with pretence. When I feign belief in God, I do not work my plans around the supposition of God's existence: I do not, for instance, plan with an eye to the Day of Judgement. All that I work into my plans is my pretence itself, not the truth of that which I am pretending.

I am led to incoherence; I have to work into my plans both the supposition that you will do it, and the supposition that you will not.¹⁶

This is not just a theoretical point. We do sometimes find ourselves in situations, situations both of mere reliance and of fully fledged trust, where we feel the pull of such conflicting states. Suppose I have been trusting you to do something. Gradually I find myself becoming convinced that you will not do it. How do I act? I find myself hatching contingency plans. Perhaps I even find myself starting to implement them. As I let myself fall, I put out my arm to save myself. What do we say in these circumstances? That I only partially believe you will let me down? That I only partially trust you? What we say will surely depend upon the details. My point is that we treat my trust in you, and my belief that you will let me down, as in conflict. The more likely we are to ascribe the one, the less likely to ascribe the other. Sometimes you can witness the conflict in yourself; at others it catches you unawares. You think that you trust your partner to stay. But suppose that they leave, and you find that you have already organized your life to cope without them. Now we surely say that your trust was less complete than you thought.

trusting vs. believing that
trust will be let down

My second point concerns the observation that whether we can bring ourselves to trust will partly depend on whether we think it right to do so: whether or not we think the trust appropriate or justified. Sometimes this will be because it is only my belief that it is appropriate to trust you that leads me to think that you are reliable. For instance, I might think it appropriate to trust you to do a certain thing only if you invite me to do so; and it might be that very invitation that leads me to think you reliable, since it signals a readiness to take on the responsibilities that my trust would bring.¹⁷ In other cases, though, I might think trust inappropriate even though I have no doubts about your reliability. I might believe that if I spilled my life's problems to you, if I put myself in your hands, you would do just the right thing. But I have not known you for long enough; we do not have the right kind of relationship. I cannot bring myself to trust you, not because I doubt you, but because I think it would not be right.

2nd constraint: we trust
only when we consider the
trust appropriate

can refrain from trusting for
reasons of appropriateness that
don't have to do w/ assurance in
the trustee

Trust and Testimony

I have argued that trust does not require belief; only reliance. Judith Baker has raised some considerations that seem to throw this into question.¹⁸ She discusses the trust that someone might have in a friend who has been accused of a crime, the trust that they might display when their friend assures them that she is innocent. She writes:

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

¹⁶Those familiar with Bratman's work on intention will see a parallel with what he says there. You can intend to do something without believing that you will succeed; but you cannot intend, whilst believing that you will fail. And again some explanation for this asymmetry lies in the need for my intentions to be consistent with my beliefs if they are to be worked into my plans. See Michael Bratman, *Intention, Plans, and Practical Reason*, (Cambridge: Harvard University Press, 1987) Ch. 3.

¹⁷Thomas Scanlon notes this possibility in 'Promises and Practices', *Philosophy and Public Affairs* 19 (1990) pp. 199-226 at p. 211.

¹⁸Judith Baker, 'Trust and Rationality', *Pacific Philosophical Quarterly* 68 (1987) pp. 1-13. Some of Baker's discussion strikes me as more appropriate to the issue of loyalty: appropriate, that is, to the question of the attitude I should adopt towards third parties regarding my friend's innocence, rather than to the attitude I should adopt towards my friend. But I shall not go into this here.

innocent. And what some outsiders find amazing is just the fact that serious belief continues in the face of rising evidence against it. As a strategy then, this response amounts to an arbitrary denial of the phenomena which raise the problem in the first place.¹⁹

Anyone who endorses an account like mine will feel the force of this passage. It is surely right that when we trust a friend, we do not simply act as if we believe what they say; we really believe them. Yet have I not characterized trust as a kind of acting-as-if? In fact I think that Baker's requirement is quite compatible with what I have said. To summarize my response in advance: when I trust my friend in the way that Baker has in mind, I trust her to speak knowledgeably and sincerely. And, because I trust her in these ways, I believe what she says. But none of this requires that I believe that she is knowledgeable or sincere.

Let me start by looking at our attitudes towards our own beliefs; in particular, towards our own future beliefs. Once again I shall start by looking not at trust but at simple reliance. Suppose I discover that in the future I will come to believe a certain proposition that I do not currently believe. What should my attitude now be to that proposition? Bas van Fraassen has shown that if there is a disparity between my current attitude towards a proposition and the attitude that I think I will have, then I incur a certain kind of vulnerability.²⁰ More precisely: suppose my conditional probability for a certain proposition p , on the supposition that I will assign p probability x , is itself other than x . Then I will be vulnerable to what van Fraassen calls a *Dutch strategy*: there will be a set of bets which will be fair given my probability assignments, but from which I will be guaranteed to emerge the loser. Of course, it is unlikely that anyone will actually offer me such a series of bets. Nevertheless my vulnerability here is indicative of a general practical vulnerability. Unless I bring my current and future beliefs into line, I run the risk that my actions will undermine each other.

Suppose then that I bring my beliefs into line in the way van Fraassen suggests: I change my current belief so that I currently assign the proposition the same probability that I believe I will later assign it. I rely on my future self to be knowledgeable. Does this mean that I must believe that my future self will be knowledgeable? Surely not. I need not believe that, and if I do not, then I could not will myself to do so. But I do not need to believe; all I need to do is rely.²¹

I can rely on my future self to be knowledgeable, without necessarily believing that I will be; and the consequence of this reliance is that I endorse my future opinions. If this happens in the first person case, why not expect it to arise in our interactions with others? I depend on information that I get from others. In many cases I adopt what they tell me as my own belief. To do this, do I have to believe that they speak knowledgeably to me? Moreover, since it is no use their being knowledgeable if they are

¹⁹'Trust and Rationality' p. 6.

²⁰Bas van Fraassen, 'Belief and the Will', *The Journal of Philosophy* 81 (1984) pp. 235-256.

²¹This is not the only way of resolving the disparity. If I do not trust my future self, I might try to ensure that I will retain my current belief; or failing that, I might at least try to ensure that I will not act on my future beliefs. Examples of such situations have been given by David Christensen in 'Clever Bookies and Coherent Beliefs', *The Philosophical Review* 100 (1991) pp. 229-47. I might, for instance, have taken a drug that will cause me to believe that I can fly. Christensen presents this example as a criticism of van Fraassen's requirement that I should change my current beliefs to bring them into line with my future ones. But the fact that there are some cases in which this is obviously the wrong strategy does not show that it always is.

lying to me, do I also have to believe that they speak sincerely?²² If so we seem to have a problem. I can surely decide to rely on a person's testimony; yet I cannot decide to believe that they speak knowledgeably and sincerely.

the difference is in one's willful ability to rely (vs. believe)

As in the first person case, I do not need to believe. Reliance is good enough. Here we have part of the answer to the problem Baker raised. When I trust my friend I rely on her to speak knowledgeably and sincerely. As a result of that reliance, I believe what she says. The belief follows from the reliance in the sense that I would be failing to act on the supposition that she speaks knowledgeably and sincerely if I did not believe what she said. As a result of deciding to rely, I come to form new beliefs: practical rationality has consequences for theoretical rationality.

But this is only part of the answer to Baker. The example she gives is an example of trust, not of mere reliance. This makes an important difference. We can see people as sources of information in two rather different ways. We can see them simply as indicators. Seen in this way, a person is like a measuring device: they respond to the environment in various ways, and we infer from their response to what the environment is like. We can, of course, rely on a measuring device even when we are not sure it is accurate; we might have nothing better to go on. But we would be plain foolish to do so if we did not check its accuracy when the occasion arose. We have reason for checking that it is accurate; and we go on relying on it in so far as we have evidence that it is.

There is a question whether we could, even in principle, always check whether people speak accurately.²³ But whatever the truth of this, there are certainly times when it does not seem right to check up on people. I think that the reason is that we are not simply relying on them to speak knowledgeably and sincerely. We are trusting them to do so; and that trust is something that we value in itself. Trust in a friend's testimony is typically a central component of the relationship that we have with them. This is what is happening in the case Baker describes. We have a reason for going on believing our friend because the trust is something we value. Seen in this way we also get some account of the attitude we have to a friend who lies to us, or who speaks authoritatively on a subject they know nothing about. The problem with such behaviour is not simply that it misleads, in the way that a faulty measuring device misleads. It is that it involves a breach of trust. I cannot defend my lying by saying that I knew you would not be misled, since even if it were so your trust would still have been betrayed.

why trust deserves slightly different treatment than reliance w.r.t. the requirement of belief

I suggest then that sometimes we can trust a friend to speak knowledgeably and sincerely, without believing that they will. And as a result of this we will believe what they say. Of course there will be a problem if one of the things they say to us is that they speak knowledgeably and sincerely. Aren't we committed to believing that? I think not: I did not say that our trust would have no bounds. Nevertheless, we need not reject the assertion. We just need to take a different stance towards it. Imagine a more likely case. Suppose a friend says: 'Believe me; I am telling the truth'. Some might think that this is a pointless thing to say. Either we believe her already, in which case it is redundant. Or

believing what they say ≠ believing that they speak sincerely

↳ avoiding the circular/self-referential case this presents

²² Whether we treat this as a new problem, not present in the first person case, will depend upon our attitude to self-deception. If you think that people can lie to themselves, then sincerity will be a requirement even in the first person case. For a discussion of the idea that we need an act of the will to overcome the scepticism engendered by the possibility of self-deception see Bas van Fraassen 'The Peculiar Effects of Love and Desire' in B. McLaughlin and A. Rorty (eds.), *Perspectives on Self-Deception* (Berkeley: University of California Press, 1988) pp. 123-156. My way of putting the point (not van Fraassen's) is that we need to trust ourselves.

²³ See C. A. J. Coady, 'Testimony and Observation', *American Philosophical Quarterly* 10 (1973) pp. 149-55; John Hardwig, 'The Role of Trust in Knowledge', *The Journal of Philosophy* 88 (1991) pp. 693-708.

we do not believe her, in which case we are certainly not going to believe this. But there is an even greater problem: surely we could not just decide to believe her statement that she is telling the truth, even if we wanted to. Nevertheless we know that just occasionally a phrase like this will make all the difference. It will be this that tips us from an attitude of disbelief to one of belief. How could it? My answer is that it is not a statement that we straightforwardly believe or disbelieve. It is a statement that invites us, or implores us, to take a stance: a stance of trust. That is something we can sometimes decide to do. Once we have done it, belief in what she says will follow.

I say much the same about the direct appeal 'Trust me'. Baier considers it, and rejects its force on parallel grounds to those we have mentioned: either we already do trust, in which case it is redundant, or, we do not, and it is powerless.²⁴ But we should not dismiss the appeal so readily; nor should we be blinded by the fact that it is often used by those we have no intention of trusting. Sometimes it does make a difference. The person who says it invites us to take a stance: the stance of trust. Sometimes that is enough to make us accept the invitation: perhaps because we had not realized it was open to us; perhaps because we thought we had already taken it but had not; perhaps because their acknowledgement that we would need to take this stance signals their acceptance of the responsibilities that would accrue to them as a result, or simply because it make the trust appropriate in a way it would not have been before.

why the 'trust me' phrase might actually have utility, contra Baier

I started this paper by claiming that we need to distinguish trust from belief, since we can choose to engage in the former but not in the latter. Now, however, it seems that I have given a kind of recipe for coming to believe something. Find someone who will tell you it is true and trust them to speak truly and sincerely; belief will follow. We can decide to follow the recipe, so surely we can decide to believe. Let me say two things in my defence. Firstly, there are, as I have stressed, limits on whom we can trust to do what. These will preclude us from employing the recipe to acquire belief in many of the things it would be useful to believe. In particular, they will preclude us from employing it to immediately believe things we currently disbelieve. One of the constraints on trusting you to do something is the lack of a belief that you will not do it. Suppose I believe that not-p, and wish I believed that p. Suppose, further, that I discover that you believe that p. Will I be able to change my belief simply by trusting you? No, since when you tell me that p, I will believe that you are not speaking knowledgeably and sincerely. So I will not be able to trust you to do so. Employing the recipe might be a way of overcoming agnosticism; it is not a direct method for up-ending existing belief. Of course, if we decide to trust someone it is likely that in time they will say things that are in conflict with our beliefs; and when they do so we might change what we believe. To say that this happens is not to say that we exploit the recipe: we did not start out trusting them just in order to change this belief. Nevertheless, we can trust someone with some expectation of the sort of direction they will lead us in, and of the kinds of revision that this will eventually bring to what we now believe. This brings me to the second thing to say in my defence. Over time we do come to acquire new beliefs as a result of our decisions. Consider the decision to do graduate work in a university very different from the one in which you were an undergraduate. That is bound to change some of your philosophical beliefs and to change them, moreover, in directions of which you have some inkling at the start. That might be part of the point of the move. Similarly with deciding to trust: in time it might bring us to revise our beliefs, in ways that are not wholly unexpected. But we know that this happens, so it is quite right that my account is able to accommodate it. This was not what I was denying

↙ caveats to the ability to come to believe, seemingly freely at will

²⁴Baier, 'Trust and Antitrust', p. 244.

when I denied that we could decide to believe. What I was denying was that we can directly and immediately decide to form a belief in the way that we can sometimes directly and immediately decide to trust someone to do something. And this, I think, is a distinction that I have respected.