

PAUL FAULKNER

A GENEALOGY OF TRUST

ABSTRACT

In trusting a speaker we adopt a credulous attitude, and this attitude is basic: it cannot be reduced to the belief that the speaker is trustworthy or reliable. However, like this belief, the attitude of trust provides a reason for accepting what a speaker says. Similarly, this reason can be good or bad; it is likewise epistemically evaluable. This paper aims to present these claims and offer a genealogical justification of them.

1.

The conversations we have with friends or with well-disposed neighbours have a different character than those we have with adversaries or pushy salespersons, but what, if anything, epistemically distinguishes these conversations? In all these examples we run the risk of acquiring a false belief if we accept what our interlocutor says and our evidence that this risk is small needn't always divide these cases; we can have evidence that what is said is false in all cases. However, while these conversations need not be distinguished in evidential terms, they are naturally distinguished in terms of trust. We trust our friends and well-disposed neighbours in a way in which we do not trust our adversaries and pushy salespersons. This trust need not be based on evidence. Indeed, there is a tension between accepting what someone says because of trust and doing so because one has evidence that what they say is generally true. Equally, trust need not be responsive to counter-evidence but can be demonstrated by continuing to accept what someone says despite the evidence to the contrary. However, if trust need be neither based on evidence nor responsive to counter-evidence, then trust is an attitude that is different from belief. If trust expresses belief, it is our believing the person rather than merely believing what that person says.¹ The initial epistemic contrast sought is then that our conversations with friends or well-disposed neighbours tend to be characterised by trust whereby we believe what is said *because* we believe them.

This paper aims to develop these ideas. In trusting a speaker we adopt a credulous attitude. This attitude is basic and cannot be reduced to the belief that a speaker is trustworthy or reliable. Nevertheless, it is not an attitude we adopt universally; rather we extend trust to some but not all speakers. Thus the

Paul Faulkner

central claims are that testimonial belief can be based on trust, where trust is a discriminating and therefore epistemically evaluable attitude. And being based on trust identifies testimony as a unique way of acquiring knowledge. The other main aim of this paper is to show how these central claims can be established by a genealogical argument. That we should accept testimony on the basis of such a basic discriminating attitude is naturalistically quite plausible and I want to argue this by extending the genealogical account Bernard Williams presents in *Truth and Truthfulness* (2002).

2.

A genealogical account of some phenomenon explains this aspect of human life by showing how it could have come about. Evolutionary accounts are genealogies: they explain particular aspects of human life in terms of the circumstances of our Pleistocene ancestors and the forces of natural selection. Moreover, all genealogies are like evolutionary accounts in that they explain aspects of human life in naturalistic terms broadly construed.² However, not all genealogies need be sensitive to the evidence in the way that evolutionary accounts must be. It is possible for a genealogy to be *imaginary* and explain some aspect of human life by showing how it follows from an *imagined State of Nature*, where this is abstractly characterised in terms of basic human limitations and needs. Such imaginary genealogies are nevertheless explanatory, Williams claims, because they offer a functional account of the phenomenon to be explained – showing that this phenomenon is necessary for human life being as it is – and do so by showing how it would be rational for individuals in the imagined State of Nature to aim for this state of affairs, were they able to appreciate it.

In *Truth and Truthfulness*, Bernard Williams offers a genealogical account of what he calls the *virtues of truth*: *Accuracy* and *Sincerity*. These are the dispositions to care about the truth of one's beliefs, and to come out with what one believes. Williams's genealogy offers an explanation of *our valuing* these dispositions or virtues of truth and in particular of our giving *intrinsic value* to Sincerity. Williams's genealogical argument for our being in this situation starts by imagining a State of Nature consisting of a primitive social group with limited technology and no writing – a description that would fit most social groups in human history. Any such society will be structured by cooperative relations and will consequently need information to be communicated between individuals. Given that an individual can only be at one place at one time, individuals will often gain what Williams calls *purely positional advantage*; that is, by virtue of their location at a time, an individual can come to possess information that would be good for another individual to share. It follows that even in the State of Nature, thus minimally characterised, Accuracy and Sincerity are desirable from the social point of view; they will be socially valued because pooled information is a social good and necessary for many cooperative endeavours. However, possessing the disposition of Sincerity need not always be

in an individual's best interest. Williams gives the example of the hunter who has found prey that he would rather keep for himself and his family. Given that the State of Nature describes something approximating an actual social group, not all relations will be cooperative, some will be competitive. This raises the problem that:

The value that attaches to any given person's having this disposition [Sincerity] seems, so far as we have gone, largely a value for other people. It may obviously be useful for an individual to have the benefits of other people's correct information, and not useful to him that they should benefit of his. So this is a classic example of the "free-rider" situation. (Williams 2002, 58)

The problem is that the collective valuing of Sincerity does not itself give an individual a reason to value Sincerity, or be sincere.

3.

It is worth developing, in a way that Williams does not, the extent of this game-theoretic problem. The basic problem is that while it is always in an audience's interest to be informed, sincerity needn't best serve a speaker's interest. A case that stereotypically illustrates this is the purchase of a used car. When a nervous buyer asks a salesman about the car's history, his purpose in doing so is to acquire information that will help him decide whether to purchase the car or not. The buyer wants to buy a car but would rather buy nothing than a wreck. However, the used car salesman's primary interest is to ensure a sale and his communication will be guided by this purpose. If the buyer recognises that the salesman would prefer to sell his worst stock, despite preferring any sale to none, then the buyer recognises that the salesman has some motive for insincerity and this makes the question of whether or not to accept the salesman's assurances the same hard question that is faced in the Prisoner's Dilemma.³ It is the same hard question because there is a symmetrical pattern of pay-offs. Credulity, like a prisoner's silence, exposes the buyer to exploitation. And like the prisoner, the buyer recognises that there exists a motive for abusing any trust. So there is the danger, as in the Prisoner's Dilemma, that there will be no trust, no sale and a bad outcome for all.

This stereotypical situation illustrates the problem of trust in communication; the depth of this problem results from the fact that interlocutors' purposes in communication can easily divide in this way. They can do so because what makes communication advantageous to speakers is that it allows them to influence audiences: in getting another to accept what one says, one can get them to think, feel and act in specific ways. On the other hand, what makes communication advantageous to audiences is that it can be a source of information about the non-immediate world.⁴ In the State of Nature, it gives access to the information speakers have through their purely positional advantage. Now suppose that the society hypothesised in the State of Nature is extensive enough for nested social

Paul Faulkner

groups and so for a speaker and audience not to be members of the same immediate social group; or, better, suppose that the society hypothesised is a loose structure of small groups sharing an environment. Given this representation of the State of Nature, two things could be true of a communicative situation. First, interlocutors can reasonably be presumed to be guided by self-interest. This is not to assume that interlocutors are selfish but simply that it is their own individual beliefs and desires, or in economic terms their preferences, which explain their communicative behaviour. Second, interlocutors might be ignorant of each other's preferences, and specifically an audience might lack information about a speaker's reasons for communicating. If this is true, the audience will not be able to assess the risk of being misinformed; in game-theoretic terms, risk becomes uncertainty because the audience cannot assess the probability of acquiring a false belief.⁵ Nevertheless, it is known that a speaker's preferences can motivate the speaker to be insincere. Thus, even though an audience might be in a position of needing to know something, since acquiring a false belief would leave the audience in a worse position than ignorance, when these things are true of the communicative situation, it would not be rational to accept a speaker's offer of information. The extent of the problem is thereby a certain scepticism of what others say – a certain scepticism concerning testimony.

4.

Though he does not develop the issue in these terms, Williams is surely aware that this is its nature. The source of the problem, he argues, is that Sincerity has only been given instrumental value; its value is given by that good which follows from knowledge being communicated, such as the good brought about by cooperative endeavour. However, when Sincerity is valued in this way it is easy for its value for an individual to be at odds with its social value. Telling what one knows is not always in an individual's best interest and those that have the disposition of Sincerity are thereby exposed to 'free-riders'. Since the State of Nature represents a basic society, this shows, Williams claims, "that no society can get by ... with a purely instrumental conception of the values of truth" (2002, 59). What any society requires is that individuals have internalised Sincerity as a disposition, where this is to say that individuals are motivated to act in a sincere way simply by the description of this way of acting as sincere. Where this is true, Sincerity will have *intrinsic value*.⁶

Williams recognises a danger here: if things are left like this – Sincerity must have intrinsic value – it can seem mysterious how this could be so. Why should the description of an act as sincere be a motivation to act in that way? But if some account of intrinsic value is offered – Sincerity must be internalised as a disposition if cooperation is to be generally secured – the account of Sincerity's value appears reductive, or instrumental. Williams's claim is that genealogy can achieve "explanation without reduction" (90). It can do so because Sincerity's

having intrinsic value can be understood in terms of two conditions. First, that it is “necessary, or nearly necessary for basic human purposes and needs” that individuals generally have this disposition, and second that their having this disposition “make sense to them from the inside, so to speak” (92). An imaginary genealogy that starts from an abstractly characterised State of Nature establishes the first condition. And the real genealogy, or cultural history, implicit in conceptual analysis can then explain how we make sense of Sincerity. In conjunction with the imaginary genealogy, this conceptual excavation is justificatory; it explains why we are right to intrinsically value Sincerity in the way that we actually do.

We understand Sincerity as having intrinsic value through its relation to other values that allow its value to be intelligible. Here “the fictional history of genealogy, which aims to bring out the necessary, structural features, is replaced by the real history of specific cultural determinations” (93). And turning to *our* social history, Williams again plausibly claims that *we* understand Sincerity as having intrinsic value through its relation to trust and our valuing trustworthy behaviour. Sincerity is trustworthiness in speech. This is more than the avoidance of lying; our being trustworthy can require our lying. Equally, it is not simply the disposition to say what one believes; one can implicate falsehoods by saying what one believes and so be untrustworthy by doing this. Williams gives the case of a speaker saying “Someone has been opening your mail” when they are the culprit. In this case the speaker expresses a true belief and so is sincere in the ordinary sense but through implicating a falsehood the speaker is not sincere in Williams’s technical sense of manifesting the disposition of Sincerity. What the disposition of Sincerity thereby requires is the ability to judge “what beliefs, and how much of one’s beliefs, one may be expected to express in a given situation” (97). This is not the kind of thing one can legislate about, so having the disposition of Sincerity does not amount to following a rule. Rather, it requires a sensitivity to the nature of the relationship one has, or seeks to have, with one’s interlocutor. Not all relationships are trusting. We have the idea that some people deserve more in the way of truth than others, which in the simplest terms is expressed in the classical distinction between ‘friends’ and ‘enemies’. A friend might expect the truth and so rely on what one says or implies, and if one is insincere or simply carelessly in error, “there is room for recriminations and a retreat from [friendly] relations” (112). The same could not be said of all relations. So we value Sincerity because and to the extent that we seek to define our relationships as trusting.

We want people to have a disposition of sincerity which is centred on sustaining and developing relations with others that involve different kinds and degrees of trust. (121)

Thus, it is through taking its place amongst the nexus of concepts and values that centre on trust that Sincerity comes to have intrinsic value for us.

5.

Williams's account of Sincerity so far is as follows. Even in the State of Nature we frequently suffer positional disadvantage and so have to rely on others' reports. Trust is then demonstrated in "the willingness of one party to rely on another to act in certain ways" (88). Since this dependence can be exploited, the pooling of information generates "game-theoretical problems about securing assurance" (89). So we value Sincerity, at least in part, because we value being able to rely on others. However, we are able to rely on others only because we have come to see value in Sincerity and for us this value is located in the connections between Sincerity, trust and our valuing trusting relationships.

At this point it needs to be recalled that Sincerity is only part of the story. The final good in the State of Nature is the pooling of information, and this also requires that speakers be accurate. When an audience trusts a speaker the implied belief that the speaker is sincere is not sufficient, on Williams's account, to make it reasonable for the audience to accept what the speaker said. Even given such belief, we can still reflect on whether to accept what another says and, Williams claims, "[t]hese reflections are appropriate just in virtue of the fact that assertions are direct expressions of belief" (79). Thus:

Assertions play their role in the transmission of knowledge just because they are taken to be direct expressions of belief *and the speakers are taken to be reliable*. (79; my emphasis)

That a speaker is believed to express a belief is not sufficient to make it reasonable for an audience to accept what is said unless it is also believed that the speaker is reliable, or possesses the disposition of Accuracy.⁷ Trust, and so cooperation in communication, therefore requires more than interlocutors believing each other to have the disposition of Sincerity; in addition each must believe that the other also has the disposition of Accuracy.

6.

With this extra condition, a fault line appears in Williams's account of how assertions play their role in the transmission of knowledge. One way of revealing this is by considering what, on Williams's account, makes trust rational. That is, what makes it rational to cooperate in communication? And the answer to this can only be whatever supplies the grounds for an audience's belief in Sincerity *and* Accuracy, or reliability. One of Williams's central insights, I think, is that the fact that interlocutors' relations are trusting can supply grounds for these beliefs. However, it does so on Williams's account only because trusting relations introduce various informal sanctions on untrustworthy behaviour. Williams's solution to the problem of trust in communication is then straightforwardly game-theoretic. We trust others not to pursue their own self-interest when it involves deceiving us because there exist sanctions on their acting in this way. These sanctions make it in

others' best interest to be trustworthy. We recognise that these sanctions exist, so we have grounds for believing others are reliable in certain ways.

The problem, as Williams is aware, is that "no line of argument which sets out from a game-theoretical formulation of the problems of trust could possibly show that trustworthiness had an intrinsic value" (90). Consequently, in so far as the solution to the problem of trust in communication is to revolve around Sincerity being taken to have intrinsic value and our making sense of this intrinsic value in terms of Sincerity being a species of trustworthiness, it must also be true that trust should be taken to have intrinsic value. However, Williams's conception of trust is quite general; it is "the willingness of one party to rely on another to act in certain ways" (88). This is a description, which, for all its interpersonal reference, could cover both our relation to other people and our relation to the inanimate world. We can trust our car to start in the morning in the sense that we demonstrate a willingness to depend on its functioning in this way. Just as we trust speakers for information, Williams claims, because we believe that they are reliable, so too we trust our car to start because we believe that it is reliable. The value of trust thus understood and placed then lies in the benefit dependence brings. In the case of communication, the value of trust will thereby be *no more* than that others can be a source of information and knowledge about the non-immediate world. But this is a return to instrumental value at odds with Williams's claims that Sincerity must be given intrinsic value and that Sincerity is trustworthiness in speech.

The resolution of the problem of trust found in the State of Nature requires that trust be intrinsically valued and this, I want to argue, requires individuating a certain conception of trust. It requires individuating a conception of trust under which trust is *sufficient* to explain an audience's acceptance of what a speaker says. That we do identify an attitude of trust that can be explanatory in this way is shown by the contrast with which this paper started. We accept what friends say because we trust them, and not because of any supporting evidence. In the same way, making it clear that acceptance is based solely on the supporting evidence is a way of distancing oneself from any implication of relationship that trust would imply. To make sense of our giving intrinsic value to sincerity we then need to explore the connections between Sincerity and this sense of trust.

A non-game-theoretic solution to the problem of trust in communication requires a shift in focus; Williams is right to claim that such a solution must be premised on Sincerity being given intrinsic value. However, this in turn requires an account of how trust can itself give an audience reason for accepting what speakers say independently of any judgement of reliability. Trust is a willingness to be dependent, so much is true; but our trusting others can involve the expectation that they be trustworthy just because we trust them to be so, and not, for instance, because this is in their interest or because there are sanctions on their acting untrustworthily. Thus, Williams's genealogy must address a dilemma, if it is to offer a non-game-theoretic solution to the problem of trust: either it is possible

Paul Faulkner

to give an account of trust which shows how the attitude of trust can itself be sufficient for acceptance; or it fails to give an account of how Sincerity, as a species of trustworthiness, is intrinsically valued.

7.

Williams is correct to say that trust is a willingness to be dependent. One reason that we can demonstrate this willingness is that we believe that our dependence will be unproblematic. Our trust then expresses the belief that the trusted is reliable in a certain respect, and trust is just this belief in the given circumstance of dependence. However, a distinction should be drawn within trust. Where relations are trusting, such that discovered insincerity would provoke “recriminations and a retreat from these relations”, our willingness to rely is not grounded by a belief about reliability but by the presumption that the trusted will respond in a certain way to our dependence. We can show a willingness to trust because we make a presumption about the trusted that reflects beliefs about the nature of the relationship had with them. To mark this distinction, let me identify *predictive* and *affective* notions of trust. ‘Predictive’ because “[w]e can trust one another to behave predictably in a sense”, Hollis notes, “which applies equally to the world at large.” (Hollis 1998, 10)⁸ ‘Affective’ because there is a purely interpersonal sense of trust where adopting an attitude of trust can make one susceptible to various reactive attitudes if the trusted proves untrustworthy. Now we give Sincerity intrinsic value, Williams claims, by its connections with trust and our valuing trusting relations. This, I would then claim, must be interpreted as referring to the relations between Sincerity and *trust in the affective sense*, and so with our valuing relations that are structured by such trust. This is because if I show a willingness to depend on your doing something simply because this is what you always do, then there need not be any room for recrimination if on this occasion you act otherwise. By contrast, there is plenty of room for recrimination if I show a willingness to depend on your doing something because I believe that my relationship with you is such that I can presume you will do this thing when you recognise my dependence on your doing so.

So one could say that trust generally involves a willingness to depend because one expects one’s dependence not to be exposed. However, where trust is predictive and so covers our relation to both other people and the inanimate world, this expectation is *merely a matter of belief in reliability*. By contrast, when trust is affective this expectation is *placed on the trusted and identified in terms of a susceptibility to certain reactive attitudes*. These reactive attitudes receive expression in a range of informal sanctions that result from “recriminations and a retreat from trusting relations”, as Williams observes. This difference between the expectations we can have in trust is further shown in the grounds we have for these expectations. When I expect my car to start, or other drivers to stick to their side of the road, my expectation is simply grounded by those observations that support my belief that

things are reliable in this respect. But in expecting it *of you* that you do something because I depend on your doing so, my expectation is not grounded by any belief in your reliability but by the belief that you can recognise my dependence and by the presumption that this dependence gives you a reason for acting in the way I expect. Let me say then:

Subject *A* predictively trusts *S* to ϕ if and only if *A* knowingly depends on *S* ϕ -ing and expects *S* to ϕ (where *A* expects this in the sense that *A* believes that *S* will ϕ).

Subject *A* affectively trusts subject *S* to ϕ if and only if *A* knowingly depends on *S* ϕ -ing and expects *S*'s knowing this to motivate *S* to ϕ (where *A* expects this in the sense that he expects it of *S* and will be susceptible to that reactive attitude characteristic of trust's being let down were *S* not to attempt to ϕ for this reason).⁹

On Williams's account, the problem of trust in communication is resolved by giving Sincerity intrinsic value. Sincerity is trustworthiness in speech. A speaker is trustworthy if he fulfils that expectation constitutive of an audience's trust. For instance, consider the case of the mail opener who says "Someone has been opening your mail". This speaker would be trustworthy in a predictive sense if their audience's expectation were merely that they reliably tell the truth. However, this speaker would be untrustworthy in an affective sense because they do not see their audience's need for information as a reason to tell what they relevantly know, and so they would not satisfy the expectation an audience would have in affective trust. Resolving the problem of trust in communication, on the present proposal, then requires an account of what makes it reasonable to hold such expectations of speakers, and of how our having these expectations is related to their satisfaction. The domain of informal sanctions is the correct place to look for such an account and so the correct place to look for a solution to the problem of trust in communication, as Williams suggested. But not because such sanctions restructure the pattern of pay-offs and so offer grounds for beliefs about reliability, but rather because our propensity to certain reactive attitudes illuminates what we bring to communication.

8.

An account of why we have expectations of people cannot be the same as an account of why we hold the reliability beliefs we do since we have a significant liberty with respect to the former that we do not have with the latter. Belief in reliability is not directly subject to the will. Certainly, we can attempt to ascertain reliability, but as Hume emphasised, any observation of constant conjunction between two objects or events will lead to the expectation that one object or event will always be accompanied by the other. And this expectation "*is more properly an act of the sensitive, than of the cognitive part of our natures*" (Hume 1740/1978, §1.4.1, 183), at least in the sense that we cannot choose to adopt it or not. Since the attitude of predictive trust is simply the expectation of reliability in a situation of dependence,

Paul Faulkner

this attitude is equally not one that we can choose to adopt or not. The same is not true of affective trust however. We can choose to trust in this sense. For instance, Holton (1994, 63) gives the example of a reformer choosing to employ someone recently discharged from prison. The reformer can trust her new employee not to steal even though the only evidence available to her suggests that he will do so. Moreover, the reformer could continue to trust her new employee, at least up to a point, even if this evidence was salient through his stealing from her; in such a situation, she could choose to give him 'one last chance'. This suggests that affective trust is an attitude that is comparable to intention in that it is an attitude that, with certain limitations, one can choose to adopt.

Nevertheless, affective trust would seem to imply belief, since there seems to be the implication that an audience believes what an affectively trusted speaker says. However, one cannot choose to believe and so this raises the question of how it can be possible to choose to trust. The answer, I think, is to deny that affectively trusting a speaker implies believing what the speaker says. Affective trust implies *acceptance*, rather than belief. This distinction is owed to Cohen (1992). In his terms, the belief that *p* "is a disposition ... to feel it true that *p* and false that not-*p*, whether or not one is willing to act, speak or reason accordingly." (4) By contrast "to accept that *p* is to have or adopt a policy of ... including that proposition ... among one's premises for deciding what to do or think in a particular context, whether or not one feels it to be true that *p*." (4) What matters to the present discussion is not the exact nature of either belief or acceptance so much as that this distinction can be made. In particular, what matters is that there is an attitude of accepting what a speaker says that involves committing oneself to treating what the speaker says as true irrespective of whether or not one also believes what the speaker says. Insofar as one can accept a proposition without believing it to be true, one can choose to affectively trust a speaker and so choose to accept what the speaker says without this implying any implausible voluntarism about belief.

A further difference between acceptance and belief that Cohen emphasises, and which fits the present concern, is that one can have good reasons for accepting a proposition that are not good reasons for believing that proposition. In the case of the reformer, if her new employee tells her that of course he hasn't stolen from her, he has reformed – or that, though he did steal the once from her, he won't steal again – the reformer has non-epistemic reasons for accepting what the ex-convict says; namely, her desire to re-integrate the ex-convict into the community and preserve her belief that people, given the chance, are trustworthy. In general, prudential reasons, moral reasons or reasons based upon friendship commonly motivate audiences to trust speakers and thereby to accept what they say. Moreover, our relations can be so structured by trust that they generate obligations to accept what is said; at least to the extent that we desire to maintain this trust and avoid provoking another's resentment. Acceptance is ordinarily sufficient to satisfy these *non-epistemic* obligations, and it is so just because it is a commitment to reason and act on the truth of what a speaker said.¹⁰

The State of Nature is fully social and so structured by relations of trust and friendship that generate non-epistemic obligations to accept any offer of information. However, it is such relations that are exploited by ‘free riders’. Thus, until some account is given of how affective trust can be *epistemically reasonable*, Williams’s demand that trust express a belief in reliability will be felt.

9.

Williams’s genealogy, I argued, faces a dilemma whose resolution requires an account of how the attitude of affective trust can suffice to explain an audience’s acceptance of testimony. However, this account cannot merely demonstrate that acceptance solely on trust can be reasonable since reasons can be moral and practical as well as epistemic. Thus the dilemma will remain unless trust provides an epistemic reason for acceptance. The worry is that any such reason must ultimately be based on a belief in reliability. However, all the while trust is conceived predictively, Sincerity and Accuracy must be coupled: a belief in accuracy does not suffice for the prediction that what a speaker says is true unless it is also believed that the speaker is sincere, or has the disposition of Sincerity. However, once trust is conceived affectively, a belief in reliability is not necessary for the possession of a reason for believing that what a speaker says is true because affective trust suffices to give an audience a reason for thinking precisely this.

In the context in which it is adopted, affective trust provides such a reason; that is, trust provides an audience *A* with a reason for thinking that *p* is true when this is what a speaker *S* says. It does so because the expectation constitutive of *A*’s trust involves the presumption that recognition of *A*’s positional disadvantage with respect to whether *p* will motivate *S* to be informative. Affective trust is then defined in terms of *A* having this expectation of *S* in a context of depending on *S* for this information. So in this context, *A* believes that *S* can recognise his informational dependence and since *S* says that *p*, *A* will presume that *S* says this because it addresses his epistemic need. However, this implied presumption is just that *S* is trustworthy: that *S* has seen *A*’s epistemic dependence as a reason to say what he knows and been moved by it. The presumption that *S* is trustworthy provides *A* with a reason to accept what *S* says just as the belief that *S* is trustworthy would.¹¹

What is important here is that ‘trustworthiness’ is not merely sincere expression of belief: the person who opens one’s mail and says “someone’s been opening your mail” is sincere but not trustworthy. ‘Trustworthiness’ is being understood as the fulfilment, by the trusted, of the expectation the truster has in trust. In the case of affective trust, this is *A*’s expectation that *S* will be moved by recognition of *A*’s positional disadvantage to tell *A* whether *p*. Thus *A* doesn’t merely trust *S* to be sincere but to further get things right; Williams is correct to observe that recrimination can be appropriate not merely if *S* is insincere but also if *S* was “careless” or even “merely wrong” (2002, 80). The presumption of trustworthiness

Paul Faulkner

that is implied by affective trust thereby includes the presumption that the speaker has got things right or is manifesting the virtue of Accuracy. The implication is that *S*'s telling is reliable. Given that *S* is addressing *A*'s informational or epistemic needs, if *S* says that *p*, then it is (objectively) probably true that *p*; or at least this is what *A* must presume. Thus affective trust provides *A* with a reason to think that *p* is true, which makes it epistemically reasonable for *A* to accept that *p*. The attitude of trust provides such a reason because it implies the presumption that *S* is trustworthy in this strong sense.

Affective trust itself is sufficient to give an audience a reason for acceptance. It does not require the support of belief in reliability. But evidence of reliability can be relevant to the reasonableness of trust since the reason trust provides is eminently 'defeasible'. The root of the reason provided by affective trust is a certain presumption and this presumption can be false and its falsity can be evident in a way that undermines any reason that the attitude of trust would otherwise provide. Trust is *not* always reasonable, and we do not act like we believe it always is. We do not, after all, trust everyone. That individuals' interests can be driven by raw self-interest and in direct competition is sometimes entirely salient; witness the perils that can confront the purchase of a second hand car. However, this prompts the question: to what extent must trust be based on or otherwise sensitive to the evidence if it is to be reasonable?

There is no simple answer to this question because the reason provided by trust is neither based on evidence nor defeated by evidence in the way of belief-based reasons. Certainly, we often do have evidence that the trusted will see our dependence as a reason to be trustworthy. But what is important is that such evidence is not essential to trust. Certainly, the evidence can undermine trust and we need to demonstrate some sensitivity to when this is the case. But what matters is that undefeated counter-evidence need not be defeating and trust can involve a wilful insensitivity to such counter-evidence. When the evidence *obliges* that one believe another untrustworthy, trust would be unreasonable. But there can be scope for reasonable trust up to this point, and whether any trust would be unreasonable – like when Sincerity requires one to reveal one's beliefs – is not something that can be specified in advance of the details.

10.

What is important here is that evidence does not have a bearing on when trust is reasonable in the same way that it has a bearing on the justification of belief. That the rationality of trust is not determined by the evidence in the same direct way as the justification of belief allows for a certain rational liberty in having one's beliefs and actions guided by trust. Moreover, this liberty is essential to creating and sustaining trusting relations since trusting relations can require that one give the *benefit of the doubt*, which is something that one would not be justified in doing if trust were only reasonable when supported by evidence.

Affective trust can create and sustain trusting relations because it involves a dependence that is salient to both parties. This dependence is exaggerated by an attitude of expectation which makes the truster further dependent on the trusted's good will or character. Were the trusted to prove untrustworthy, the truster would not be merely let down, they would, in some way, also be betrayed. Since the truster's dependence is salient, the truster will believe the expectation they tie to this dependence will equally be so, and with it the presumption this expectation involves, namely that relations are trusting. When this is mutually recognised, it can come to form the background to a conversation, and interlocutors' expectations can be attuned to this background. The expectations thus formed are not just a matter of subjective probability: interlocutors *expect things of* each other. Such expectations, I've claimed, can be identified in terms of our being susceptible to certain reactive attitudes, and this accords with Anscombe's observation that:

It is an insult and it may be an injury not to be believed. At least it is an insult if one is oneself made aware of the refusal, and it maybe an injury if others are. (1979, 150)

That is, just as untrustworthiness in a speaker can prompt certain reactive attitudes in an audience, so an audience's *failure to extend trust* can prompt comparable reactive attitudes in a speaker. It would do so where there is a presumption that relations are trusting.

The expectation that trust will be extended is the expectation of acceptance not belief. It is an insult and it may be an injury to have an audience reject what one says, but what ordinarily matters here is the audience's refusal to accept rather than any private mental attitude. The insult to a speaker lies in an audience's not treating an offer of information as true; it would not be caused by the audience's private reservations, since there is mutual knowledge that even the best intentioned speakers can go wrong. The most that can be expected is reliability. The bond created by trust is not that interlocutors should think the same thing, but that they should share their knowledge.¹² Where there is a background of trust, each interlocutor can rightly presume that the other will see their positional disadvantage as a reason to be informative, and, more generally, see their dependence as a reason for action. What trusting relations hereby establish, Hollis suggests, is *local identities* whereby each gains a reason to act for the good of the 'team' (1998, 137ff). Trust is a way of creating and sustaining local identities, and to reason in terms of such identities is to be trustworthy.

11.

The fact that affective trust can constitute a local group's identity can then explain how information can be pooled in the State of Nature. What is needed is some account of how an individual *A* can know that *p*, and tell *B* that *p*, who in turn can tell *C* that *p*, who can then know that *p*. Williams's account, as described, involves the thin notion of predictive trust with its requirement that each interlocutor

Paul Faulkner

believes the other to be reliable. This account, I argued, is at odds with the need to give Sincerity intrinsic value, where this is required if we are to solve the problem of trust and not to generate a certain scepticism concerning testimony.

However, suppose that interlocutors *A*, *B*, and *C* correctly presume that their relations to one another are trusting, and do so because they identify themselves as a group, which in the State of Nature can but be local. In this case *B* will depend on *A* for information as to whether *p*, and *C* will similarly rely on *B*, because each will make certain presumptions about how the other will view their epistemic dependence. Since these presumptions turn out to be true it is not merely that *B* has a reason to believe *A*, and *C* has a reason to believe *B*, so each has a reason to accept that *p*, but it is further that *A*'s knowledge that *p* has been shared and is now collective knowledge, or pooled information.

The idea that it is knowledge that *B* and *C* acquire could be variously understood. For instance, suppose knowledge is understood in terms of the existence of a reliable connection. Given that *A* knows that *p* and intentionally addresses *B*'s epistemic need to know whether *p*, *A*'s telling that *p* is reliably connected to the truth of *p*: if *p* were false, *A* wouldn't believe that *p* so wouldn't tell *B* that *p*. The same can be said of *B*'s telling *C* that *p*. Now suppose that knowledge requires knowledge-supporting evidence. Since *B* and *C* lack a reliability belief for the testimony they accept, they lack knowledge supporting evidence. However, when *B* accepts *A*'s telling that *p*, *B* presumes that *A* can address his informational or epistemic needs; on this account, this is a need for evidence, which is something that *A* does possess. Something similar could then be said about *B*'s testimony to *C* except *B* possesses the evidence only by belonging to the local group – the 'team'.

However, what is crucial for any epistemological explanation of *B* and *C* acquiring knowledge are two claims. First, that *B*'s and *C*'s acceptance is reasonable based as it is on affective trust, which is a reason-giving state. Second, that the presumptions that *B* and *C* make in trust are true. It is this second fact that delivers a reliable connection or supporting evidence. And this second fact is in turn delivered by the prior existence of *A*, *B* and *C* as a local 'team'. It is thereby group identity that secures the truth of the presumptions that *B* and *C* make in trust, where the truth of these presumptions then legitimises the idea that trust is a means of transmitting knowledge or creating collective knowledge.

12.

The transmission of knowledge in the State of Nature, which defines *our* genealogy, thereby presupposes a certain kind of relationship: a relationship of affective trust. Since these relationships are also partly constitutive of local groups, this is to say that we share knowledge with some people but not all. Equally, we extend trust to some people but not to all. This could be put schematically: *we extend trust to, and share knowledge with, some but not all X*, where 'X' refers to an individual, group

or entity that it is possible to identify and trust. One central task of this paper has then been to offer a genealogical justification of our having this practice. The justification is that this practice is needed for cooperation, which the pooling of information is an instance of and a prerequisite for. In terms of the State of Nature, we might extend trust to the family, the broader kin group or members of the local society that could be called the tribe. We do not extend trust to strangers.¹³ Moving beyond the State of Nature, the principle that we extend trust to some but not all X still holds but now ‘members of the local society’ are more numerous; a characteristic of our modern world is that we presume a certain relationship, ‘fellow citizen’ for instance, with an enlarged group of persons and can show a willingness to extend trust accordingly. However, it still remains true that it is facts about our relations with the trusted that determine whether or not trust is objectively reasonable or knowledge-giving. And this is compatible with our relations changing and extending trust being one way of changing our relations. Again this is Williams’s point that Sincerity cannot be reduced to a rule like “Only tell the truth to X”.

This account of the rationality of affective trust provides a solution to the basic problem of trust outlined above. It does not provide a solution in the sense that it offers a strategy for resolving problems of distrust or resolving the problem of trust conceived sceptically, but because it is part of a philosophical explanation as to why such distrust is not pervasive. This explanation is genealogical. The basic problem of trust found in the State of Nature determines that every society must find some way of recognising the intrinsic rationality of trying to get one’s offer of information right and accepting others’ offers of information. In our society, this rationality is determined by the nature of affective trust. This solves the basic informational problem because it involves an audience presuming that a speaker will see his epistemic dependence as a reason for being informative, in making this presumption the audience gains a reason to trust the speaker, and when things go well the speaker’s knowledge is thereby pooled. There need be no problem of trust because interlocutors, through identifying themselves and others in terms of social groups, presume the existence of trusting relations. This solution to the problem of trust is expressly non-game-theoretic because the rationality which governs trusting relations cannot be wholly captured in terms of the gilding of raw self-interest. This is to agree with Hollis that interlocutors’ trust-based reasons “are not sentiments, which can be represented in utilities, but normative features of a situation which they have partly inherited and partly created” (1998, 141).¹⁴

REFERENCES

- Anscombe, G. E. M.** 1979. “What Is It to Believe Someone?” In C. Delaney (ed.), *Rationality and Religious Belief*, pp. 141–51. London: University of Notre Dame Press.
- Boyd, R. and P. J. Richerson.** 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Paul Faulkner

- Cohen, L. J.** 1992. *An Essay on Acceptance and Belief*. Oxford: Clarendon Press.
- Dunbar, R., C. Knight, and Camilla Power** (eds). 1999. *The Evolution of Culture: An Interdisciplinary View*. Edinburgh: Edinburgh University Press.
- Faulkner, P.** 2007a. "On Telling and Trusting." *Mind* 116(464): 1–28.
- Faulkner, P.** 2007b. "What is Wrong With Lying?" *Philosophy and Phenomenological Research* 75(3): 535–58.
- Hardin, R.** 2006. *Trust*. Cambridge: Polity.
- Heinrich, J., R. Boyd, et al.** 2001. "In Search of Homo Economicus: Experiments in 15 Small-Scale Societies." *American Economic Review* 91: 73–8.
- Hollis, M.** 1998. *Trust Within Reason*. Cambridge: Cambridge University Press.
- Holton, R.** 1994. "Deciding to Trust, Coming to Believe." *Australasian Journal of Philosophy* 72(1): 63–76.
- Hume, D.** 1740/1978. *A Treatise of Human Nature*. Oxford: Clarendon Press.
- Johnson, A. and T. Earle.** 1985. *The Evolution of Human Societies: From Foraging Group to Agrarian State*. Stanford: Stanford University Press.
- Levinson, S. C. and P. Jaisson** (eds.). 2002. *Evolution and Culture*. Cambridge, MA: MIT Press.
- Luce, R. D. and H. Raiffa.** 1957. *Games and Decisions*. Mineola, NY: Dover Publications.
- Moran, R.** 2006. "Getting Told and Being Believed". In J. Lackey and E. Sosa (eds.), *The Epistemology of Testimony*, pp. 272–306. Oxford: Clarendon Press.
- Richerson, P. J. and R. Boyd.** 2000. "Complex Societies: the Evolutionary Dynamics of a Crude Superorganism." *Human Nature* 10: 253–89.
- Sperber, D.** 2001. "An Evolutionary Perspective on Testimony and Argumentation." *Philosophical Topics* 29(1&2): 401–13.
- Welbourne, M.** 1981. "The Community of Knowledge." *The Philosophical Quarterly* 31: 302–14.
- Williams, B.** 2002. *Truth and Truthfulness*. Princeton: Princeton University Press.

NOTES

- 1 The significance of this distinction is also argued for by Moran, see his (2006).
- 2 Moreover, there is a large literature concerned with giving an evolutionary account of cooperation. For instance, see Levinson and Jaisson (2002), Boyd and Richerson (1985) and Dunbar et al. (1999)
- 3 Though this coordination game differs in that it is one-way. See Hardin (2006, 46ff).
- 4 Sperber offers a genealogical (evolutionary) argument to the conclusion that communication has evolved and stabilised only because it serves *both* these purposes. See Sperber (2001).
- 5 See Luce and Raiffa (1957, 13).
- 6 That is, "Sincerity will have *intrinsic value* for us". Williams is interested in our *giving intrinsic value* to Sincerity rather than it *being intrinsically valuable*; if it were the latter, we would have an external reason to be Sincere. Thus, Williams states: "[I]t is right only to consider human attitudes. Indeed, it is part of the naturalistic outlook of this inquiry that it should be seen as an exercise in human self-understanding." (2002, 60).

- 7 Or again: "It may be said that a hearer never has a reason for believing that P which lies just in the fact that a given speaker has told him that P. He has to believe also that the speaker (on such matters, and so on) is a reliable informant." (77-8).
- 8 My two senses of trust compare to Hollis's 'predictive' and 'normative' senses.
- 9 I offer these definitions of trust in Faulkner (2007a,b).
- 10 'Ordinarily' because on occasions belief might be required; for instance, if the speaker were an innocent friend protesting his innocence despite the mounting evidence to the contrary.
- 11 This is worked out in much more detail in Faulkner (2007a).
- 12 Compare Welbourne (1981).
- 13 There is empirical evidence of significant differences in willingness to trust across cultures. See Johnson and Earle (1985) and Heinrich et al. (2001). However, there is also evidence that a willingness to cooperate with strangers, in some respects, is extensive. See Richerson and Boyd (2000).
- 14 Thanks to Miranda Fricker, Ted Hinchman, Mike Martin, Matt Soteriou, Bob Stern and audiences in Lampeter, Madrid and at the fourth annual *Episteme* conference at Rutgers University.

Paul Faulkner completed his Ph.D. at University College London in 1999, worked as a temporary lecturer at UCL for two years and since then has been a permanent lecturer at the University of Sheffield. To date he has largely worked on the epistemology of testimony and trust.