The Metaphysics of Trust

Abstract:

I argue against the claim that the fundamental form of trust is a 2-place relation of A trusting B and in favour of the fundamental form being a 4-place relation of A, by ψ-ing, trusting B to φ. I characterize trusting behaviour as behaviour that knowingly makes one reliant on someone doing what they are supposed to do in the collaborative enterprise that the trusting behaviour belongs to. I explain how trust is involved in the following collaborative enterprises: knowledge transfer – i.e. telling someone something; maintaining a relationship; and passing responsibility for an action on to someone else. And I finish by showing how our talk of trust in non-collaborative contexts – e.g. trusting a branch to support one’s weight – may be explained by reference to the central sort of collaborative trust.

Key words: collaboration; reliance; communication; Faulkner; Simpson; Jones.

1. *Introduction*

I develop a rather straightforward account of what trust is. We may think of trust in the first instance as an attitude – a dispositional state – that is manifested in trusting behaviour. I argue that trusting behaviour is characterized by the fact that the truster is taking themselves to be working together with the trusted one on a single action or enterprise. The trusting behaviour belongs to such a collaborative enterprise and at the same time makes the truster reliant on the trusted one to do what they are supposed to do in the very same enterprise. While plenty of philosophical treatments of trust link it with reliance and with collaboration, this way of putting them together is new, and, as I will indicate, surprisingly powerful as an explanation of ethically significant examples of trust.

I will begin in section 2 by clarifying the metaphysical relationship between trust and reliance. I argue that we should base an account of trust on an account of trusting behaviour, and focus on the 4-place relation expressed as follows: ‘A, in φ-ing, is trusting B to ψ’. In opposition to some recent arguments that the fundamental form of trust is the 2-place relation: ‘A trusts B’, I argue that this 4-place relation is fundamental in building an account of what the attitude of trust is.

You might behave towards someone in a way that makes you reliant on them without that behaviour being a manifestation of trust. Annette Baier (1986, 235) introduces the example of the citizens of Königsberg who allegedly would set their clocks by Immanuel Kant’s highly regular habit of walking round the town at the same time every day. In setting their clocks to his walks they were making themselves reliant on the regularity of his behaviour, but it would be wrong to say that they were trusting him. I propose that what makes the difference between this sort of case and genuinely trusting behaviour is simply that in cases of genuine trust φ and ψ belong to the same enterprise. What determines what the trusted is supposed to do is the activity that the truster themselves is engaged in. Kant and the Königsberg citizens were not collaborating in a common task; but, if they had been, then the citizens, in setting their clocks to Kant’s walks, would indeed have been trusting him.

I develop this notion of collaborative trust in section 3, explaining how it applies to ethically significant examples of trust, like maintaining a relationship, trusting someone to be telling you the truth or trusting someone to manage some task on their own. We employ a weaker conception of trust when we talk about trusting a clock to show the right time, trusting your instincts in working out what to do, trusting a dog not to scratch the furniture or trusting a branch to hold your weight. Actual collaboration is not involved in this second category of cases, but in section 4 I show that, in common with cases of collaborative trust, these cases involve doing something within some enterprise or activity that makes one reliant on the other thing doing what it is supposed to be doing within that same enterprise or activity. In some cases, like trusting a branch to hold one’s weight when climbing a tree, the trusted thing is not literally supposed to do what you trust it to do. But I suggest that we are willing to extend our talk of trust to the branch to exactly the same extent that we are willing to extend our talk of there being something it is supposed to do to the branch.

A possible payoff of this account is that it may help us direct research into social sources of distrust – for example, distrust of scientific experts, politicians, journalists, etc. Instead of focusing only on epistemic inadequacies in the situation of the distrusters we should also consider the implications of a failure of collaborative engagement on their part. If you are not engaged in collaborative enterprises with scientists, politicians, journalists, etc., you won’t be trusting them. Social scientists, who think of themselves as engaged in such collaborative enterprises, may struggle to understand the mindsets of anti-vaxxers or conspiracy theorists who aren’t. They may assume that the only solution is to communicate better to these deluded people, somehow persuading them of the trustworthiness of experts. But more productive perhaps would be an approach that analyses the sort of collaborative enterprises engaged in by distrusters, including, for example, the enterprise of maintaining a status in social media, and thinking about what it would take for such enterprises to become genuinely knowledge-seeking collaborative enterprises with a role for experts as well as consumers.

1. *Trust and reliance*

When we say that A *trusts* B we are usually describing A as having some kind of attitude towards B. But when we use the progressive form of the verb and say that A *is trusting* B we are not simply describing A’s attitude towards B, but describing something about the way A is behaving in relation to B. You might trust your climbing partner, and, in particular trust them to keep hold of the rope you are attached to, even when you are not acting on that trust. You might have trusted them for years. But when you deliberately let go of your handhold and have nothing supporting you except the rope, you are, in so doing, *trusting* them to keep hold of your rope.

The attitude of trust and trusting behaviour are bound up with one another in an obvious way. This is because the attitude is a dispositional state of mind characteristically manifested in trusting behaviour. Your trust may remain dormant until triggered by encountering a situation that calls for it to be manifested in trusting behaviour. Now, in providing a philosophical account of a dispositional state there are two ways to go. One is to provide a reductive account of the attitude in terms of other attitudes. For example, we might take trust to be a special kind of belief or confidence – perhaps the belief in someone’s trustworthiness or perhaps a combination of belief and some emotional attitude. The other - non-reductive - way to go is to account for the dispositional state in terms of its characteristic manifestations. In this case it is to provide an understanding of what trust is by providing an understanding of what trusting behaviour is.

I will pursue the second of these approaches. There seems to be something distinctive about trusting behaviour that may be appealed to in characterising it as trusting behaviour. Then the attitude of trust may be characterized as the disposition to behave in that sort of trusting way in certain circumstances. This is at odds with accounts, like those of Paul Faulkner (2015) or Jacobo Domenicucci and Richard Holton (2017), that take the attitude of trust to be primary, with trusting behaviour understood in terms of its relationship with that attitude. A quick argument against these accounts is that trusting behaviour may be quite independent of the attitude of trust and is compatible with the attitude of distrust. Suppose I lack the attitude of trust towards my climbing partner and in particular towards them keeping hold of the rope I am dangling from. I may not be at all disposed to behave in a way that trusts them to hold onto the rope. Yet at the same time I may have no option when trying to extricate myself from a tricky situation than to release my handhold, trusting that they will keep hold of the rope while I find another handhold. I am trusting them to keep hold of the rope even though I do not trust them to keep hold of the rope. My attitude might be one of distrust, despite my behaviour being that of trusting someone, and this means that what makes the behaviour trusting is not that it derives from an attitude of trust.

The distinctive feature of trusting behaviour has something to do with reliance. In letting go of my handhold and trusting my climbing partner to keep hold of the rope I am making myself reliant on my climbing partner to keep hold of the rope. And this is something I might do even when full of distrust. Observing that we can rely on things without trusting them, most approaches tend to treat trust as reliance plus X, and concentrate on debating what should take the place of X. But I think we should step more slowly here. For one thing, when we consider the metaphysics carefully it becomes clear that the attitude of trust is not actually a form of reliance at all. Jacobo Domenicucci and Richard Holton (2017, 151-2) point out that even when trust and reliance go together the attitude of trust *explains* the reliance and so should not be identified with it. I might lend someone my car because I trust them not to damage it, but I don’t lend them my car because I rely on them not to damage it. My relying on them not to damage it comes as a result of my lending it to them, not before it, whereas my attitude of trust comes before the lending.

Saying that A is relying on B is to describe a current state of reliance, which is not itself an attitude or mental disposition. What it means to say that I *am currently relying* on my climbing partner to keep me safe is roughly that I am in a state in which it is necessary for the realization of some positive result in me that she does what needs to be done to keep me safe. So reliance is a modal state. What it means for A to Ψ, relying on B to φ, is for A to Ψ in order to achieve some goal because they are working on the assumption that B will φ, where the successful achievement of that goal depends on the assumption not being mistaken. For example, in letting go of my handhold I was relying on my climbing partner to keep hold of the rope. And this means that in letting go of the handhold I was aiming for some goal which included not being badly hurt; and in so doing I put myself in a position in which it was necessary for the realization of that goal that my climbing partner keep hold of the rope.[[1]](#footnote-1)

The attitude of trust is not itself the modal state of reliance. Rather, it is manifested in trusting behaviour that puts the agent into a modal state of reliance. But even this idea that trusting behaviour always makes one reliant has been challenged recently by several authors. Jason D’Cruz writes: “One may trust another person but have no need to rely on them to perform an action. (For example one may discover that the necessary action has already been performed by someone else.)” (2018, 241) And Christopher Thompson (2017, 649) and Harold Noonan (2021) in the same vein provide the example of someone offering to buy the drinks for their friend’s party not knowing that the drinks have already been bought. Not to undermine your friend’s confidence, you let your friend buy the drinks as well and indeed trust them to buy them, but do not rely on this, as you already have the drinks for the party.

I’m not convinced by this example. Even though you’ve already got enough drink for the party it might still be the case that in some sense you are relying on your friend to buy the drinks. Perhaps you are trying not merely to make sure that you have enough drink for the party, but also to have your friend supply the drink. You need them to be doing it because you care about boosting their confidence. And the successful achievement of this goal of having your friend buy the drinks depends on them doing their part. If we were convinced that there is no sense in which you are relying on your friend to get the drinks, I think we might then question whether it is clearly right to say that you are trusting your friend to get the drinks.

Domenicucci and Holton (2017) present a different argument for separating trust from reliance.

In trusting, we grant discretion, whether to act, or to judge, or even just to feel. Granting such discretion is not all there is to trust; indeed it may not even be necessary. But it is central. (2017, 151)

The idea is that lending someone your car may count as a manifestation of trust in them just in virtue of the fact that you are giving them some power. You are allowing them to do what they will with your car. But although this sounds plausible initially, it can’t be exactly right. Unless you are relying on them to use that power well your act wasn’t a manifestation of trust. If you lend them your car granting them discretion over its use without having any interest in what happens to it, it would be a mistake to describe that action as a manifestation of trust. Simply handing over your car to someone is not the same as entrusting your car to someone, even though they both involve granting discretion to that person. So while acts of trust may be described as acts of granting discretion, they must also be acts that put one into a state of reliance.

Domenicucci and Holton point out that you may trust someone without trusting them to do anything in particular, and this might suggest that trust does not give rise to reliance, since reliance seems to be tied to particular things that you are relying on.[[2]](#footnote-2) But this does not seem at all clear to me. In lending someone my car I may not be merely trusting them to drive at a certain speed or not to crash it into a wall. But I am trusting them to take care of the responsibility I am giving them. And I am to some extent reliant on their doing so.

Consider saying that you trust your teenage child. This is quite vague and may mean something like: ‘Within a contextually determined range of situations in which I would rely on her to act in a certain way I would make myself reliant on her to do so.’ For one sort of parent that range might include situations where, if the child were given the freedom to decide what to do for herself, she might or might not act responsibly. The parent, in manifesting their trust of their child, is reliant on her acting responsibly in a standard range of situations where the need to act responsibly arises. They are reliant on her to do so because they still have as one of their own enterprises in life the task of ensuring the safety and wellbeing of their child. It is necessary for the achievement of the parent’s goal here that the child acts responsibly in that contextually determined range of situations.

However, the child may be looking for a deeper sort of trust – a trust that would be manifested by giving up on the enterprise of ensuring the safety and wellbeing of the child and handing over complete responsibility to them. With this sort of trust, the child would not in fact be letting the parent down by acting irresponsibly. But even here there is some reliance that must result from the parent’s trusting behaviour, because there is a very big difference between washing your hands of your child and trustingly letting them go their own way to learn from their own mistakes in life. When you let your child go in this sense you are only trusting them to make use of their freedom and your trust will only be let down if they don’t do that.

So I am rejecting the claim that trust may be manifested in behaviour that requires no corresponding behaviour from the trusted person. The appearance that this might be so is due to the vagueness in what counts as the range of such requirements in any instance of trust and also in the possibility that what may be required of the trusted person might be as general as that they make good use of the freedom they are being granted. This claim that I am rejecting has led Domenicucci and Holton (2017) and also Faulkner (2015) to argue that fundamentally trust is a two-place relation and not a 3-place relation. While reliance is taken to be a 3-place relation – ‘A is relying on B to φ’ - they argue that the basic form of trust is that of A trusting B, not that of A trusting B to φ. They treat trust as an interpersonal attitude like love. Although there is a 3-place relation of loving someone for something, this may be understood in terms of the 2-place relation of loving someone, and not vice-versa. In the same way, they argue that what it is to trust someone to do something may be understood in terms of an understanding of what it is to trust someone, but not vice-versa. Indeed Faulkner goes further in arguing that the 1-place property of being trusting is even more fundamental.

I have argued that they are wrong to think that trust may be understood except in terms of trusting behaviour that involves relying on someone else’s behaviour. Indeed, I am proposing that the most fundamental form of trust is the 4-place relation: A, in ψ-ing, trusts B to φ. Only by making visible this 4-place relation when thinking about trust do we see what is the central thing that characterizes trust and explains why it would be a mistake to describe Kant’s neighbours as relying on him to make his walks around Königsberg at precise times. For we can see that there are two potential actions in this sort of trust – φ and ψ.

In the rest of the paper I develop the claim that what characterizes trust fundamentally is that φ and ψ belong to the same activity or enterprise. In cases of trusting behaviour the truster assumes that they are working together with the trusted – in a collaboration of some sort. Kant and his neighbours were not working together. But to consider a different example, the person who is coming out of the bushes waving a white flag and the enemy who is deigning not to shoot them *are* collaborating. It takes two to surrender – someone to give themselves up and someone to accept their surrender. And when you do emerge waving a white flag you are doing your part in this collaborative activity trusting the other to do what they are supposed to do in the same activity.

1. *Collaboration*

The precise claim that I will develop and defend is that a subject, A, in Ψ-ing, is trusting B to ϕ, if and only if A, in Ψ-ing, is doing what they are supposed to be doing as part of a collaborative enterprise, relying on B doing what they are supposed to be doing – namely to be ϕ-ing – as part of the same enterprise. You are trusting your collaborator if, in doing your part, you make yourself reliant on the other to do their part. Given what I have said about reliance, this means that you are doing your part in order to achieve some goal because you are working on the assumption that the collaborator will do their part, where the successful achievement of that goal depends on the assumption not being mistaken.

If two people are working quite independently of one another towards the same end, then neither collaboration nor trust is involved. For example, two of us might share the ambition of getting a common friend a nice birthday present. We both act with that goal in mind. But although the two of us may be doing the same kind of thing, there are two separate actions we are embarked on. I am doing what *I can* to get our friend a nice birthday present. The other one is doing what *they can* to get our friend a nice birthday present, and these are different particular enterprises. To the extent that our activities are independent of those of the other we are not relying on each other or actually collaborating, and trust does not come into it.

It is an important question, and one that I won’t try to answer here, what exactly is it that distinguishes working together from working independently towards the same goal. Within the literature on joint intentionality a central dividing issue is whether it is possible to provide an analysis of what it is to work together within a single enterprise in terms of what each individual is doing. But the history of attempts to reduce joint intentionality to individual intentionality is not encouraging, and I think some moves in the philosophy of trust fit into this history. For example, Karen Jones (1996) and Paul Faulkner (2007) have included in their accounts of trust a condition that looks like an attempt to analyse the idea of two people working together on the same enterprise in terms of the expectations and actions of each. Jones makes it a condition of our trusting someone that we have “the expectation that the one trusted will be directly and favourably moved by the thought that we are counting on her.” (1996, 4) Faulkner picks up on this condition as follows:

A trusts S to φ (in the affective sense) if and only if

1. A knowingly depends on S φ-ing and
2. A expects S’s knowing that A depends on S φ-ing to motivate S to φ. (2007, 882)

This idea that the trusted person is being relied on to be motivated by their knowledge that the truster is depending on them looks like an approximation to the idea that the trusted person is being relied on to be collaborating with the truster, and to that extent I think they are on the right track with this. But, as it stands, it is both too strong and too weak a condition. It is too strong for two reasons. First of all, we can be collaborating with people who don’t know us and by the same token we can also trust people who don’t know us. I develop examples of this later. The second thing that makes Faulkner’s condition (though not Jones’) too strong is that *knowledge* is too strong a requirement. After making an arrangement to meet someone in exactly ten years’ time at a particular place I might go to the meeting relying on them to do their part, while expecting them to have only the faintest hope that I will be relying on them by going to that meeting. I trust them to remember to go to the meeting, but I don’t expect them to know that I am depending on them.

The condition is also too weak to capture the 4-place relation of A, in ψ-ing, trusting B to φ. An example that often crops up in the literature alongside that of Kant and his neighbours is that of a burglar who relies on a householder to be out at a certain time to walk the dog or pick up their children from school. The burglar, in breaking into the house at that time, is not trusting the householder to be out – or only in the very weak sense that I consider in the last section. A version of this example I think serves to defeat Jones’ and Faulkner’s condition. Suppose that it is the burglar’s own children who the householder collects from school at this time. As it happens the burglar is relying on the householder collecting the children, and the householder knows this and is motivated by this knowledge. But on this occasion it is not the shared enterprise of making sure the children are collected from school that the burglar is engaged in; it is the completely different enterprise of robbing the house, one that they are certainly not collaborating in. Although it is true that in not being at the school gate themselves το pick up the children, the burglar is trusting the householder to be there, it is not true, or only in that weak sense, that in breaking into the house at that time the burglar is trusting the householder to be at the school gate.

What about forced collaboration? Suppose I am forced by the gang that is holding my daughter hostage to work with them on some nefarious scheme to build a nuclear bomb. If I don’t help, they will kill my daughter. I don’t want our collaborative activity to succeed and it is certainly wrong to say that I am trusting them to do their part in the building of this bomb. But I think this may be explained by the fact that the task I am engaged in is not after all the collaborative task of building a nuclear bomb. After all that is not my goal. My goal is actually to give the impression of collaborating in this task; as such it is not an example of genuine collaboration.

Note the example changes if my daughter only gets released if we succeed with the task; then I really am trusting the others to do their part. This illustrates the fact that there is no requirement that collaborations involve anything like equality. A collaborator with an oppressive regime may be acting out of fear but they are still working with it. A servant is collaborating in the master’s or mistress’s tasks. A child, learning something from a grown-up, or a citizen, learning something from an expert, is on the weak side of a power relationship, but nevertheless is collaborating with them in the joint enterprise of knowledge transfer.

Collaboration has many forms, the simplest of which is teamwork. You might be relying on your teammate to pass the ball through to you as you make a run around the back of the defence or be relying on your climbing partner to keep hold of the rope you are dangling from. In these cases you are trusting someone to do their part in some collaborative activity you are engaged in. When you trust someone to be telling you the truth you are in the same way trusting them to do their part in a collaborative activity with you. This may be less obvious, but it takes two to tell someone something. The teller has to say what they know and the hearer has to accept this as true.

Lack of trust in this case occurs when you don’t accept what the other person is saying even though you take them to be supposed to be saying what they know as part of a collaborative activity of communication. By not accepting their word you are not doing what you are supposed to be doing within this collaborative activity. That is why lack of trust is sometimes a failing, and at its worst a betrayal of a relationship.[[3]](#footnote-3) At the very least, assuming that you are accepting that they are communicating what they know, you are letting down the other person who is saying what they know. That is because they are saying to you what they know on the assumption that you are going to do your part and accept it; in saying this to you they are trusting you to accept what they say.[[4]](#footnote-4)

Indeed this feature of trust in the case of communication carries over to other cases of trust on the collaborative model. The footballer who makes the pass on the assumption that you are doing your part and running in behind the defence trusts you to be making the run in exactly the way that you trust them to be making the pass. But in other cases there may be an asymmetry in collaboration where one person acts on the assumption that the other does their part but the other doing their part is not done on the assumption that you do yours. The other person may not be aware of you relying on them, in which case they certainly won’t be trusting you to act on that reliance. For example, a student may log in to a live online lecture I am providing on the assumption that I am doing my part in the collaborative activity of my lecturing them. They trust me to show up. But I may be unaware of which students are logging in, perhaps only trusting that at least someone is out there listening. So I can do my part in a collaborative activity in which a particular student is engaged even though I am not relying on that particular student to be engaged. I am not trusting that particular student to show up.

A similar but more interesting example of this sort of lack of symmetry occurs when you accept something that someone has written or posted. They may be long dead but you are still engaged in a collaborative activity with them, where you are working on the assumption that they did what they were supposed to be doing as part of that activity. So it is possible to trust someone to have done something in the past. The long dead writer was trusting that someone would accept what they were saying, but not that any particular person would.

Collaborations require omissions as well as actions. In robbing a bank the gang has to work together, trusting each other to do their parts. But after the robbery the collaboration continues as they work together to preserve the ignorance of the authorities. This aspect of their collaboration involves *not* boasting publicly about the robbery and *not* talking to the police even if caught. Each trusts the other not to snitch on them. In general, maintaining a relationship, whether it be a relationship of honour among thieves, a friendship or a love affair requires both acts and omissions on all sides. The parties have to work at these relationships and part of that work is done by not doing things - not betraying secrets or developing love affairs with other people. And in working at one of these relationships the parties trust one another to be doing their part too.

Having a friendship involves treating your friend in some respects differently from the way you treat others. And this difference is in part a normative difference. On the one hand, friendship may involve enjoying things together and being on the same wavelength in certain sorts of conversation. But these are things that connect me to the stranger sitting next to me at a football match. What characterizes a friendship is also that you look out for one another, that you are there for each other in some respect, that you have some concern for them. And this involves something like a commitment. The precise nature of this commitment will vary depending on the nature of the friendship. If your friend fails to be there for you in the relevant respect they have let you down. That is because in making your own commitment to the friendship you are relying on them to do the same. The joint enterprise of having and maintaining a friendship requires something from both parties.

As we have seen, the joint enterprise may rely on little more than not doing certain things that one would naturally do with respect to people outside of the relationship. If I am particularly sensitive to not revealing something of myself to non-friends and some non-friend is responsible for revealing it to the wider world this isn’t a breach of trust but merely a humiliation to me. But if my friend does it, then their failure to live up to the joint commitments of the relationship undermines that relationship which I have been maintaining by satisfying my side of the deal.

This applies more obviously to certain sorts of love affairs that involve exclusivity. In committing to the relationship my partner may be relying on my equivalent commitment. In relying on me to do my part in this joint enterprise they are trusting me. But it also applies, though less obviously, to relationships that complete strangers may have towards one another. While passing someone on a lonely mountain path I smile and say hello and they look away and do not respond. I have exposed myself to being cut by them. I was relying on them to do their part in a normatively structured social relationship of politeness. The joint enterprise in this case is the creation and maintenance of this very short-term relationship. And as, by saying hello and smiling, I was relying on them to do what they were supposed to do in this joint enterprise that I took us both to be in, this just means that I was trusting them in this respect.

Note that this is slightly different from the idea that I am trusting the stranger not to turn around and murder me. I do rely on this, of course, and indeed we could say that it is a default feature of my plan as I walk through the countryside that nobody is going to murder me. But even though I am relying on others to do what they are supposed to do as far as this plan is concerned, these others are not literally collaborating with me in this plan when they don’t murder me. They may be completely oblivious to such a plan. That is why I think it sounds a bit stretched to say that in walking past a stranger I am trusting them not to murder me, whereas it sounds fine to say that in greeting a stranger I am trusting them to return the greeting.

Omissions that count as part of collaboration may be on the part of the truster as well as on the part of the trusted. For example, you may decide *not* to do something, relying on someone else to do it instead. The two of you have the same goal of getting this thing done, but you may be working on the assumption that the other one will do it, and so you go away and do something else instead. This is a clear example of trusting someone. Your reasoning is something like the following: G is to be achieved; she is going to achieve G without your help; so you don’t need to do anything about it. What distinguishes you from someone who does not have G as their goal is how you would act if you weren’t working on the assumption that the other person is going to achieve G without your help. You would conclude that you better do something about it, or at the very least worry about it.

In this sort of case you are working together on a shared enterprise. The enterprise is to get G achieved. The person you trust may have no idea that they are involved in a collaboration; as far as they are concerned they are just doing it by themselves. But even though you two are not achieving G together, you are working together on the goal of getting it done. Her part is actually to do it, while your part is to leave her to do it, and these two parts are not independent of one another. While her part is independent of you doing yours, yours is not independent of hers; your contribution depends on your assumption about what she is going to do. And you are partly responsible for getting it done even though you are not at all responsible for doing it.

In the case considered earlier of the parent letting the child go forward in life unhampered by their influence, the parent’s letting go is an act of trust. The shared enterprise is that of getting the child on to their own path in life. In some circumstances your role as parent in this enterprise would be to intervene. But at some point your part is simply to let go of your control and hand over responsibility. In doing so, you are relying on the child to make use of the freedom and get themselves on to their life path in their own way.

1. *Working with things without collaborating with them*

We often talk of trust where there is no scope for collaboration. But I think we can see why such talk is appropriate by revealing what such cases have in common with what I take to be the central notion of collaborative trust. To support this I will consider a series of increasingly strained ways to use the word ‘trust’.

1. Trusting your tools

You might trust your watch to show you the correct time or your car not to break down on the motorway. You cannot be said to be collaborating with these inanimate objects. But we do talk of your working with them. Your watch is *supposed to* show the correct time; that is its function within the enterprise you have of working out what time it is. It has a role in this enterprise. The same sort of thing may be said of your car in your enterprise of getting to some destination.

Working with tools has something in common with working with a person in a collaborative enterprise. In both cases there is something the trusted thing or person is supposed to be doing, where this normative demand on it or them is rooted in the success conditions of the truster’s enterprise. So there is a weaker conception of trust than the one I have been considering in this paper. The notion of collaborative trust I have been developing is a special case of this more general idea of working with something or somebody on an enterprise and doing something as part of that enterprise which makes you reliant on the thing or person you are working with doing what they are supposed to be doing as part of that same enterprise.

1. Trusting yourself

Suppose you wonder whether to trust yourself not to burst out laughing in a funeral. Perhaps in the end you decide to go, trusting yourself not to burst out laughing. One way to make sense of this talk of trust is to interpret it as splitting yourself into a current and a future self. In deciding to go to the funeral you are relying on yourself to do what you would be supposed to do according to the enterprise of paying your respects, which deciding to go to the funeral now would be part of. So in a sense you are working with yourself and relying on yourself; and in this sense you are also trusting yourself.

1. Trusting your dog not to scratch the furniture

In leaving my dog alone in my friend’s house I am trusting the dog not to scratch the furniture. It’s absurd to think that I am taking the dog actually to be collaborating with me in this respect. But I am relying on the dog to do what I take it to be supposed to do as far as my enterprise of respecting my friend’s house is concerned. Now it may be literally wrong to say that there is something a dog is supposed to do; it’s a creature of habit and instinct and not sensitive to normative considerations. But if that is right it is equally wrong to talk about trusting the dog or to talk of the dog letting me down or betraying my trust in it when after all it does scratch the furniture.

1. Trusting a branch to support your weight

It makes sense to ask whether you trust a branch not to break when you put your weight on it while climbing a tree. But not only is the branch not collaborating with you, but also it does not have as its function to support human beings who are climbing trees. However there is a rather strained sense in which it makes sense to say that the branch is supposed to bear your weight. Your plan to climb the tree makes use of various things, including the branch, and relies on them behaving in certain ways. These are the ways these things are supposed to behave according to your plan. The branch has a role in your plan – to bear your weight - and when you climb on it, relying it to do what it is supposed to do according to your plan, you are trusting it to bear your weight.

I think this usage should be considered to be a metaphorical extension of our more proper talk of trusting things. It is not always appropriate to extend our talk of trust to behaviour that makes one reliant on elements of one’s plan working properly. In the case of Kant’s neighbours, it may seem wrong, even metaphorically, for me to say on the day he oversleeps that he was supposed to be on time. I can imagine a case where one of his neighbours might say that they were trusting him. For example, they might in a rather petulant way say: “But he was supposed to be on time; I was banking on that.” And if they allowed themselves that petulant thought I think they would equally allow themselves to say that they were trusting him to be on time when they set their clocks. This kind of extension of the idea of collaborative trust might in theory be extended to any case of mere reliance. But our reluctance to make this extension exactly mirrors our reluctance to extend the idea of something doing what it is supposed to do in an enterprise. Suppose I knew that it was very wrong for Kant to be on time with his walk today as he really needed a bit of extra sleep for the good of his health, but still I relied on his being on time when I set my clock. When he does oversleep on this occasion, I don’t allow myself the petulant thought that he was supposed to be on time, but think instead that he has done the right thing at last. I don’t think I would extend the notion of trust to this case in these circumstances. I would resist saying that I was trusting him to be on time when I set my clock, even though I was relying on him to be so.

1. I trust you are well

Not everyone appreciates someone saying to them that they trust that they are well. It’s a rather odd thing to say. First of all it is almost always a disguised question. The conventional response is to say something like: “Yes, thank you.” It seems to be a way of asking the question about whether someone is well while not really allowing a negative answer. So I propose the following paraphrase: “May I be disposed to act in a way that relies on your health doing what it is supposed to be doing as far as my enterprise of not being concerned about you goes?”

1. *Final Remarks*

Philosophical work on trust in the last thirty years has produced many different accounts and counterexamples to just about all of these accounts, and a general sense that the target notion is moving around uncontrollably between these accounts. So it is common these days to avoid making claims about what trust is or what we mean by trust, and instead to talk about a cluster of trust concepts. Depending on what use we are trying to put the idea of trust to we might focus on one or another of these concepts. This might suggest that instead of conceptual analysis our task is to provide a genealogical account of trust.[[5]](#footnote-5) But I have been arguing that the account of trust as an attitude characteristically manifested in collaborative behaviour that makes one reliant on the behaviour of another within the same collaborative enterprise has some more fundamental status, and may be used to explain much of our talk of trust.

At the same time it might be argued that this account does not capture the most interesting aspects of trust. Very often there is some emotional aspect to trust. You might feel betrayed if you are let down when you trust someone. Saying, “But I trusted you!” has power as a moral appeal to someone – aimed at making them feel guilty. Trust seems often to be an aspect of certain sorts of relationship involving commitment and loyalty. And it can have the function of enabling co-operative behaviour when self-interest – as in the Prisoners’ Dilemma – will lead to worse results.[[6]](#footnote-6)

But the fact that these interesting aspects of trust aren’t always present shows that it would be wrong to try to shoehorn them into the account of what trust actually is. It is much better to see them as aspects of particular sorts of collaborative enterprise within which trust may have a role. For example, the possibility of betrayal seems to be associated with relationships involving loyalty. Whereas you might feel let down when someone who has committed to write a paper for a publication you are editing fails to do what you were relying on them to do, you only feel betrayed by someone if their commitment was to a joint enterprise requiring loyalty for its success. For many sorts of relationship maintaining the relationship has this kind of requirement. Loyalty involves a sort of exclusive or preferential commitment. For example, if the renegade colleague did write the paper they were supposed to write for you but sent it to a rival publication instead, that failure to do what they were trusted to do might then count as a betrayal.

References

Baier, A. 1986, “Trust and Antitrust”, *Ethics* 96(2), 231-260.

Baker, J. 1987, “Trust and Rationality”, *Pacific Philosophical Quarterly* 68, 1-13.

D’Cruz, J. 2018, “Trust within limits”, *International Journal of Philosophical Studies* 26(2), 240-250.

Domenicucci, J. and Holton, R. 2017, “Trust as a two-place relation”, in P. Faulkner and T. Simpson (eds.) *The Philosophy of Trust*, Oxford University Press, 149-160.

Faulkner, P. 2007, “On telling and trusting”, *Mind* 116(464), 875-902.

Faulkner, P. 2015, “The attitude of trust is basic”, *Analysis* 75(3), 424-429.

Hardin, R. 1996, “Trustworthiness”, *Ethics* 107(1), 26-42.

Jones, K. 1996, “Trust as an affective attitude”, *Ethics* 107(1), 4-25.

Noonan, H. 2021, “No trust is hybrid: reply to Faulkner”, *Philosophia*.

Simpson, T. 2012, “What is Trust?”, *Pacific Philosophical Quarterly* 93, 550-569.

Thompson, C. 2017, “Trust without reliance”, *Ethical Theory and Moral Practice* 20(3), 643-655.

1. See Thompson (2017) for a defence of this sort of account of reliance. [↑](#footnote-ref-1)
2. Paul Faulkner (2015) also argues that trust is not limited to a particular sort of behaviour that one is trusting someone to do, though see Noonan (2021) for what I take to be an adequate rebuttal of Faulkner’s arguments. [↑](#footnote-ref-2)
3. See Judith Baker (1987). [↑](#footnote-ref-3)
4. This is why unjustly refusing to accept what someone is trying to tell you may be described as a kind of silencing; without having to stop their mouth you are stopping someone from doing what they are entitled to do – to tell you something. [↑](#footnote-ref-4)
5. Simpson 2012 is a nice example. [↑](#footnote-ref-5)
6. Russell Hardin 1996 is an example of someone arguing for the moral importance of trust in this way. [↑](#footnote-ref-6)