



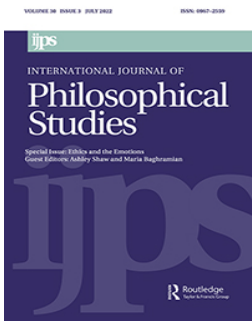
Title	Betrayal, Trust and Loyalty
Authors(s)	Stout, Rowland
Publication date	2022-09-28
Publication information	Stout, Rowland. "Betrayal, Trust and Loyalty." Taylor & Francis, September 28, 2022. https://doi.org/10.1080/09672559.2022.2121892 .
Publisher	Taylor & Francis
Item record/more information	http://hdl.handle.net/10197/25514
Publisher's version (DOI)	10.1080/09672559.2022.2121892

Downloaded 2026-05-02 00:26:37

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)



© Some rights reserved. For more information



Betrayal, Trust and Loyalty

Rowland Stout

To cite this article: Rowland Stout (2022) Betrayal, Trust and Loyalty, International Journal of Philosophical Studies, 30:3, 339-356, DOI: [10.1080/09672559.2022.2121892](https://doi.org/10.1080/09672559.2022.2121892)

To link to this article: <https://doi.org/10.1080/09672559.2022.2121892>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 28 Sep 2022.



Submit your article to this journal [↗](#)



Article views: 6647



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Betrayal, Trust and Loyalty

Rowland Stout

School of Philosophy, University College Dublin, Dublin, Ireland

ABSTRACT

I argue that while every betrayal is a breach of trust, not every breach of trust is a betrayal. I defend a conception of trust as primarily a feature of behaviour (i.e. trusting behaviour) and only secondarily a feature of a mental attitude. So it is possible to have the attitude of distrust towards someone while still trusting them in the way you behave. This makes sense of the possibility of Judas Iscariot breaching Jesus' trust, and so betraying him, even though Jesus presumably knew that Judas would do just that. This conception of trust may be spelt out in terms of making oneself reliant on somebody in a collaboration with them. Such collaborations include relationships like friendships and love affairs, as well as political activities or defending one's country against aggression. I argue that only when these collaborations involve a commitment to loyalty is a breach of trust a betrayal. And loyalty is a feature of those collaborations or relationships that exclude others – us/them collaborations.

KEYWORDS Trust; loyalty; betrayal; reliance

1. Introduction

Baier (1986) is the innocent progenitor of a tendency in the philosophy of trust to treat all breaches of trust as betrayals. She is often taken to be responsible for the idea that the difference between trust and mere reliance is that trust, when breached, leads to betrayal or the feeling of betrayal.¹ And since then, many philosophers have argued that it is distinctive of trust that it makes you vulnerable to betrayal.² But what Baier actually says is that 'trusting can be betrayed, or at least let down, and not just disappointed.' (1986, 235). This distinction between being betrayed and merely being let down is completely ignored in most of the literature, which correctly treats being merely disappointed as something else again.

I will argue in this paper that, while betrayal is always a breach of trust, not every breach of trust is a betrayal or is capable of leading to the feeling of

CONTACT Rowland Stout  Rowland.stout@ucd.ie

¹See for example McCleod's (2021) entry in the *Stanford Encyclopedia*. But I could have come up with a dozen examples.

²For example, Holton (1994) argues that 'when you trust someone to do something, ... you have a readiness to feel betrayal should it be disappointed, and gratitude should it be upheld.'

betrayal. When your trust is breached you will be let down and may feel let down, but you will not always be betrayed. This means that the relationship between betrayal and trust is not completely straightforward. I will argue that the way to make sense of the fact that only some kinds of trust result in betrayal when breached is to identify trust in general with making oneself reliant within a collaboration and to say that betrayal only belongs to certain sorts of collaborations – in particular, those that depend on loyalty.

I claim that making and sustaining a relationship, like friendship or a romantic attachment, is a collaboration of a sort. There are plenty of collaborations (and relationships) that don't involve loyalty, and where trust without vulnerability to betrayal is possible. But for some, like monogamous sexual relationships, loyalty is essential, and the failure to do what you are trusted to do, when this is to remain loyal, counts as betrayal. This means that betrayal should be understood in the following way:

X betrays Y with respect to a relationship between them that is partly constituted by the commitment to be loyal when the trust that Y has in including X in the relationship is breached by X failing to maintain this loyalty.

The rest of the paper is taken up with explaining and defending this account.

2. The relationship between trust and betrayal

I start by claiming that betrayal is always a breach of trust. This is often taken for granted in the literature on trust, but there are some problem cases that are worth looking at in this section. One problem case would be the apparent possibility of betraying someone who already doesn't trust you. The other cases that I will consider concern betraying something other than a person – for example a country, political party or set of ideals.

Judas Iscariot betrayed Jesus, we are told, by pointing him out to the authorities with a kiss. But we are also told that Jesus knew he was going to be betrayed, and we can assume that he knew it would be Judas who would do it. Given this, Jesus did not trust Judas, but he was still betrayed by him.

To respond to this I need to introduce a distinction that I will develop properly in the next section – that is between trust taken to be a feature of one's attitude towards someone and trust taken to be a feature of the way one is behaving towards someone. I will argue that it is possible to do something, trusting someone else to do something, while at the same time not have the general attitude of trusting that person. For example, if you have nothing else to grab hold of apart from the hand of your fellow climber, who you don't actually trust to keep a grip, you may grab hold of their hand trusting them not to drop you.

Jesus, we may suppose, made Judas Iscariot an apostle, allowing him to stay close to him with the other apostles. In doing so he was relying on him to

remain loyal and not to do things like point him out to the Sanhedrin. In making him an apostle and keeping him as an apostle Jesus was trusting Judas to play a part in furthering the religious cause and not to give the enemies of this cause the opportunity to damage it. He was trusting him in this way even though he was expecting him to breach that trust.

Of course, it sounds paradoxical to say that you are trusting someone even though you don't trust them. But, with the distinction between trusting behaviour and trust as an attitude in place, it is a perfectly reasonable possibility. The attitude of trust is a dispositional state; you may trust someone even when you are not manifesting that trust in any way. For example, you still have the attitude even when you are fast asleep. In the next section I will argue that the way to explain what the attitude of trust is is to explain what it is a disposition to do; and this is to behave in a trusting way. Having the attitude of trusting someone to do something is to be disposed in a contextually determined range of situations to behave in a way that trusts them to do that. You might not be disposed to behave in a trusting way towards someone in that range of situations, yet still behave in a trusting way towards them in some particular situation. And this represents the possibility of doing something, trusting the other to behave in a certain way, while not having the attitude of trusting them to behave in that way.

Soldiers might surrender to the enemy trusting them not to shoot them, while being pretty sure that the enemy will in fact shoot them. The soldiers might not trust these enemy soldiers not to shoot surrendering combatants in general, but in emerging with their hands raised they are trusting the enemy soldiers not to shoot them. Their decision on this occasion is to trust the enemy even though their attitude towards the enemy is one of distrust. If they had a reasonable alternative they wouldn't behave in this trusting way, but on this occasion they have no option but to trust the enemy.

The Judas case is more extreme, we might suppose, assuming that Jesus actually knew, rather than merely suspected, that he was to be betrayed by Judas. But this knowledge does not rule out the possibility of trusting behaviour. You might start a collaborative project with someone knowing for sure that the project will fail because they won't do their part of it. You continue with your part – behaviour that only makes sense as part of a collaborative project. The success of the project depends on the other doing their part. You continue because you don't want to be the person who is responsible for the failure of the project. In doing your part you are trusting them to do their part, knowing they won't. The project will fail and you know it will, but it is important that it is the other person who will be responsible for this.

The notion of therapeutic trust may be relevant here. Victoria McGeer introduces it in the following way:

Think here of the oft cited example of parents deciding to trust their teenagers with the house or family car, believing that their offspring may well abuse their trust, but hoping by such trust to elicit in the fullness of time, more responsible and responsive trustworthy behaviour. (McGeer 2008, 241)

In most of the literature on therapeutic trust the question is how one might have an attitude of trust even when there is no good reason to believe it won't be breached. And there may be cases where it is the attitude that is relevant. Certainly, in Baker's (1987) famous example of trusting your friend who is accused of being a spy, what your friend needs is your attitude of trust and belief, and not any particular behaviour. But in the McGeer example just quoted and indeed in most of the examples in the literature on therapeutic trust it is actually trusting behaviour that is at issue. The parents lend their house or car, trusting the child to look after it, while believing their child won't look after it. They act in a way that is trusting but their attitude is distrust.

The second sort of problem case I want to consider is betraying something other than a person – for example, your country, political party, religion, or set of ideals. None of these betrayed things are the sorts of things that might be said to trust anyone, since only people trust. But in each case there are people involved, and my claim is that any betrayal is a breach of their trust. Each of these betrayed things is constituted by a collaborative project of some sort. People who associate themselves with a country, political party, religion or set of ideals are working together to promote or defend these things. And, in doing so, they rely on others in the collaboration to do their part in promoting or defending the collaboration. As a member of one of these collaborations you are trusted by the other members to remain loyal to it. If you have no trusted role in your country or party, then there does not seem to be anything to betray.

It helps to get the logical form of betrayal clear. When a person A commits a betrayal there is something they have committed to, which they have failed to honour. And I want also to claim that there is a person or people to whom they have made this commitment which they have failed to honour. And when we describe someone as betraying someone or something we might equally be taking the object of the betrayal to be the person to whom they have made the commitment or the thing that they have committed themselves to.

Let us apply this to a problem case. Think of someone who is a citizen of a country, but has always been hostile to its values and ambitions. They are not part of a patriotic collaboration of any sort, and make no pretence of this. Now suppose they start working against the interests of the country and in the interests of a foreign power, but not using any aspect of their status as a citizen of that country to gain this advantage for the foreign power. Have they committed a betrayal of any sort? Certainly, they might be accused of

treason and of having betrayed their country. But I think this is because any citizen is automatically assumed to be part of a collaboration to foster national interests in virtue of their status as citizens. And, as part of that patriotic collaboration, they are relied upon and trusted to do their part. If the person can successfully disavow this commitment to the collaborative task of fostering national interests, then they don't seem to have really betrayed anything or anyone. The issue then is not whether betrayal depends on a breach of trust, but whether being a citizen implicitly entails involvement in a patriotic collaboration, which can't be disavowed.

Someone might respond that something like this might work for betraying a country, or a political party or football team, but not for betraying a set of ideals, as we might have allegiance to a set of ideals in isolation from anyone else. This would be a problem for my claim that all betrayals are breaches of trust. Suppose that, despite not having any political allegiance, you were a committed socialist in your own value system, and you failed to live up to this ideal on an occasion. You might be described as betraying your ideals, yet no third party is involved here.

One way to respond to this sort of example is to say that there is someone who is betrayed – namely the person themselves. We might say that the person trusted themselves and betrayed themselves. This way of talking treats the individual as two people – an earlier self who is betrayed and a later self who does the betraying. In itself, it is not a counterexample to the claim that every betrayal is a breach of trust.

It is also worth observing that use of the term 'betray' can move quite far from the core case where betrayal does require a breach of trust. For example, someone who accidentally reveals some fact about themselves through some unconscious piece of body language or Freudian slip is said to have betrayed this fact. But no real betrayal is involved. Similarly, we might say that no real betrayal is involved when we talk of someone betraying their ideals.

Suppose that instead of not living up to your ideals, you simply change them. You now espouse conservative values. Again, we might think of this as betraying your earlier self. Or we might simply reject the idea that this involves any actual betrayal if there is nobody else involved in the story. Of course, you might feel like a traitor. But this might be explained by saying that you might *feel* like a traitor without actually betraying anything or anybody. You have imagined that you are part of a group of socialists, perhaps including your parents; it's a fantasy, perhaps quite unconscious, that has motivated you in your socialist behaviour. And now in changing your ideals, it feels as if you are betraying these people, even though, of course, you are not. You might imagine your parents rolling in their graves if they knew of your change of allegiance, but since they are in fact dead you are not really betraying anyone.

I grant that intuitions are not clear in some of these examples, and there is no hard and fast principle concerning when it is appropriate to talk of betrayal. But, at any rate, the claim that betrayal involves a breach of trust looks like a reasonable starting point in this discussion.

What would be unreasonable would be the converse claim that every breach of trust is a betrayal. It is a claim made explicitly or held implicitly by many philosophers in this area. For example, McLeod (2021), while conceding that there may be some exceptions to this claim involving therapeutic trust, regards these as outliers. Now, as it happens, therapeutic trust need not be trust that is not vulnerable to betrayal. The Judas and Jesus example considered earlier might, perhaps with some modification (we might have to assume that Judas, instead of hanging himself, goes on to live a better life), be taken to be a case of therapeutic trust. But it still involves betrayal. And the actual examples of breaches of trust that don't involve betrayal are not outliers at all.

Consider someone failing to do what they are supposed to do when they are working with a team. A goalkeeper trusts their defender to stop the ball at the near post when they cover the rest of the goal, or the midfielder trusts their forward to make a run behind the defence when they pass the ball through the channel. When the trusted player fails to do what they were supposed to do and trusted to do, the other player may feel let down, but not betrayed. When my climbing partner drops the rope that I am hanging from, my feelings as I fall won't include a sense of betrayal. When I ask a stranger what the time is and it turns out that they give me the wrong answer my trust has been misplaced but not betrayed. If I greet someone on the street and they ignore me I trusted them to return my friendly greeting and feel let down and humiliated, but not betrayed. When someone breaks a promise to me – e.g. not to buy drugs with the money I give them – my trust is breached, but again I haven't been betrayed. When a journal editor has an agreement with a potential author that they should submit a paper by a certain date and they fail to do what they were trusted to do, a sense of betrayal would be entirely inappropriate. The emotional responses that would be appropriate might be resentment, disappointment and annoyance; but a feeling of betrayal is something quite different.

Now there are two possible justifications for the philosophers who take every breach of trust to be a betrayal. They may be thinking that these cases of trust without risk of betrayal are not the interesting cases – perhaps not trust in the full sense. But, in fact, these *are* interesting cases of trust, and include the currently most discussed sort of trust – namely trust in experts. Or, they may be aiming to distinguish those failures of what you rely on that merely lead to disappointment from those that lead to a feeling of being let down. And they may not take the difference between feeling let down and

feeling betrayed to be significant. By contrast, I want to mark this difference and explain it.

Think of that journal editor again. Suppose that their author, instead of simply failing to come up with the promised paper, writes it but sends it to a rival journal. Now the editor may feel betrayed. That failure of the author to do what they were trusted to do might count as a betrayal, not because of the mere breakdown of trust, but instead because the joint enterprise that they were committed to required some sort of loyalty for its success. Loyalty involves a sort of exclusive or preferential commitment, and when someone's trust in this sort of commitment is disappointed then they may feel betrayed.

This is what I explore in section 4. But first I propose a conceptual account of what trust in general is that makes space for the possibility that some instances of trust involve vulnerability to betrayal, and some do not. The sort of account of trust that is automatically ruled out by this requirement is one that simply identifies trusting with making oneself vulnerable to betrayal or the feeling of betrayal. Also ruled out automatically are those approaches that identify trust with reliance plus X, where X is an affective component involving vulnerability to betrayal. Not only are such accounts mistaken for the reasons I have been outlining, but they leave the idea of betrayal hanging unexplained.

What we need is a philosophical understanding of what betrayal is. I have argued that it is not merely to be understood as a breach of trust, but nor is it helpful to say that a betrayal is a breach of trust plus some other independent element. Certain kinds of trust when breached always involve betrayal and other kinds of trust when breached never involve betrayal. And what this means is that the difference is internal to the sort of trust that is in play, not a difference in some other factor. So my proposal is that by understanding what trust is and then identifying the special sort of trust that makes betrayal possible we get an understanding of what betrayal is.

3. Trust as Reliance within Collaboration

The account of trust I propose is that trusting someone is making oneself reliant on them within a joint enterprise. Under the category of joint enterprise I will include maintaining relationships like friendships and love affairs. The key thing is that they are collaborations. It is a commonplace observation that trust has an important role in collaborative behaviour. Annette Baier observed that 'it seems fairly obvious that any form of cooperative activity, including the division of labor, requires the cooperators to trust one another to do their bit' (1986, 232). More recently, Simpson (2012) has made the stronger claim that our notion of trust arises from our need to rely on others in co-operative activity. I am going to argue for the still stronger claim that collaborative reliance is not merely an important way

that trust figures in our lives; collaborative trust is what trust is, at least as far as the core conception of trust goes. When trusting someone you are doing what you are supposed to be doing within some collaborative activity on the assumption that the person you are trusting is doing what they are supposed to be doing within the same activity.

The first task is to defend the view introduced in the previous section that trust may be understood in terms of doing something that involves trusting someone else to be doing something, and not just as a special attitude that might be characterized without reference to such trusting and trusted behaviour. 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*. Reversing the earlier example, you might trust your climbing partner, and, in particular, trust them to keep hold of you if you are grabbing on to them, 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 partner's grip, you are, in so doing, *trusting* them to keep hold of you.

As I pointed out before, the attitude of trust is a dispositional state. 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. A classic example of this is the attempt to explain knowledge as belief plus some extra conditions. In the current discussion, 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 Faulkner (2015) or Domenicucci and Holton (2017), that take the attitude of trust to be primary, with trusting behaviour understood in terms of its relationship with that attitude.

An obvious problem with these other accounts is that trusting behaviour may be quite independent of the attitude of trust and is compatible with the attitude of distrust. My earlier examples, including that of the climber lacking the attitude of trust towards their climbing partner but still trusting them in grabbing their hand, show this. They climber does not have the attitude of

trust towards their partner nor towards the partner keeping hold of their hand. They may not be at all disposed to behave in a way that trusts the partner to hold onto them. Yet at the same time they may have no option when trying to extricate themselves from a tricky situation than to release their handhold on the rock and grab their partner's hand, trusting that their partner will keep hold of them while they find another handhold. As I argued, their attitude might be one of distrust, despite their 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 alternative approach is to say that what makes an attitude an attitude of trust is that it disposes one to trusting behaviour.

This approach depends on being able to explain what trusting behaviour is. As with most approaches, starting with Baier (1986) I assume that the distinctive feature of trusting behaviour has something to do with reliance. In grabbing their partner's hand the climber is making themselves reliant on their partner to keep hold of them. And this is something they might do even when full of distrust. But reliance by itself is not sufficient for that. As remarked by Baier (1986), the citizens of Königsberg may have relied on Immanuel Kant to take his walks around the town at exactly the same time each day when they set their clocks by his appearances, but it would be a mistake to say that they trusted him to do so.

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 my hand. 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 my hand.³

The 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 Thompson (2017) and Noonan (2021) in the same vein provide the example of a friend offering to buy the drinks for your party not knowing that the drinks have already been bought. Not to undermine your friend's confidence, you let them buy the drinks as

³See Thompson (2017) for a defence of this sort of account of reliance.

well and indeed trust them to buy them, but do not rely on this, as you already have the drinks for the party. What we have been describing as therapeutic trust is taken not to require reliance.

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. For this reason you embark on a collaborative mission with your friend of *their buying drink for the party*. And the successful achievement of this goal of having your friend buy the drinks depends on them doing their part. If we were really convinced that there was no sense in which you were relying on your friend to get the drinks, I think we might then question whether it is clearly right to say that you were trusting them 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 were relying on them to use that power well your act wouldn't count as 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.⁴ 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 at least trusting them to

⁴Paul 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.

take care of the responsibility I am giving them. And I am to some extent reliant on their doing so.

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. So 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 . And seeing that there are two actions involved in any attribution of trusting opens up the possibility that the relationship between these actions is crucial.

I 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 with their hands in the air and waving a white flag and the enemy who is deigning not to shoot them *are* working together; in a sense, they 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.

The precise claim that I 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. 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.

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 even 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. It is to relationships of this sort that I now turn, as these are the ones where the possibility of betrayal becomes relevant.

4. Loyalty

So far I have argued that betrayal is always a breach of trust, that trust is reliance within a collaboration and that not every breach of trust is a betrayal. The next claim I want to make is that those breaches of trust that do constitute betrayals are betrayals of loyalty. The collaborations that may lead to betrayal are those that involve loyalty. I will argue further that loyalty always involves exclusion. In being loyal to one person or group you have a relationship with that person or group that requires behaviour that excludes certain other people or groups.

Considering untrustworthy bank robbers can help us focus on the sort of breach of trust that constitutes betrayal. If the gang were relying on the

getaway driver to steal a fast car and bring it to the heist, and she fails to do it, they are let down by her, but not betrayed. However, if she tells the police about the plan then that does constitute a betrayal. Trust is involved in both cases. She was trusted to be there with the fast car and she was trusted not to tell the authorities, and in both cases she failed to do what she was trusted to do with the result that the heist failed. But only in the second case does this failure constitute a betrayal.

The difference seems to be that in the second case, but not the first, the failure is a breach of loyalty. And this carries over to all cases of betrayal. Judas Iscariot, in being appointed as an apostle was trusted to be loyal; the treacherous lover in being accepted into a monogamous sexual relationship was being trusted to remain loyal to it; the double agent in being given access to the secrets of one side was being trusted to remain loyal to that side; the friend of the celebrity who sells their secrets to the press was trusted in being made privy to these secrets to be loyal to the friend. So our task is to work out what loyalty is and what it is to trust someone to be loyal.

Suppose the bank robbers meet a passer-by on the way to the heist and get a commitment from them not to tell the police. The passer-by assures them that they themselves have no interest in helping the police and that they can be trusted to stay quiet. Then they tell the police, breaching that trust and destroying the success of the heist. This does not count as a betrayal. The difference between this and the betrayal of the getaway driver who also tells the police is that the getaway driver is a member of the gang. She has a relationship with the rest of the gang that the passer-by does not have, and this relationship is a special relationship partly constituted by the expectation of loyalty.⁵

It seems clear that loyalty involves partiality in some way. Loyalty is something you have towards some people and not others. But there is room for debate over whether this partiality must be understood in terms of actively excluding others from something. George Fletcher is someone who has argued that it must be.

There are always three parties, A, B, and C, in a matrix of loyalty. A can be loyal to B only if there is a third party C (another lover, an enemy nation, a hostile company) who stands as a potential competitor to B, the object of loyalty. The competitor is always lurking in the wings, rejected for the time being, but always tempting, always seductive. (Fletcher 1993, 8)

And Simon Keller claims that 'loyalty is the attitude and associated pattern of conduct that is constituted by an individual's taking something's side' (2007, 21) But Kleinig (2022) in the *Stanford Encyclopedia* entry on loyalty suggests

⁵See Zac Cogley's account of betrayal in which he argues that feelings of betrayal must be responses to the 'rupture of an important relationship' (2012, 40). However, Cogley does not characterize the importance of the relationship in terms of loyalty.

that the following examples represent the possibility of loyalty without exclusion.

Sometimes, however, the loyal friend will simply manifest the loyalty by being responsive to *B*'s need at some inconvenience. The loyal *A* will get up at 2.00am to fetch *B* when *B*'s car has broken down or will agree to be best man at *B*'s wedding even though it will involve a long flight and great expense. No third party is involved, but there will be a cost to *A*.

The loyal friend in Kleinig's example is certainly partial; he wouldn't do these things for just anybody. But he is not taking sides; the relationship that he is respecting is not one that requires being us against them. So the example has some initial plausibility. However, one might argue that although a loyal friend might behave in this way, they would not be manifesting their loyalty by doing so, but rather be manifesting their love of their friend. Suppose that, despite being willing to go out of their way for their friend they are unable, due to some sort of cowardice, ever to take their friend's side against others. They'll do anything for their friend unless opposition with others is involved. I think that this would mean that their claim to loyalty would be undermined. In any case, I will use Fletcher's conception of loyalty, as that is the one that is relevant when considering betrayal. Failing to get out of bed at 2 a.m. to collect your friend is no kind of betrayal. But failing to stand up for your friend when they are being attacked or insulted is.

Some relationships essentially involve a commitment to loyalty in Fletcher's sense. I include here not just personal relationships like friendship and monogamous romantic attachments, but also some relationships that someone may have with people who are in the same group as them. The idea here is that groups and relationships may be normatively constituted. What makes us members of such a group (and equivalently what it is for us to have the defining relationship for that group) might be normative commitments and entitlements. Let me go through some examples, starting with friendship and loving relationships.

Having a friendship involves treating your friend in some ways differently from how 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 commitment to preserve some exclusivity of knowledge within the relationship undermines a key feature of that relationship which I have been maintaining by satisfying my side of the deal. It is a betrayal, because that key feature is loyalty.

There are many kinds of friendship and many kinds of loving relationship, and not all involve loyalty and not all involve the possibility of betrayal. You might have a relationship that does not require loyalty in the sense of being on each other's side – potentially against others. One sort of friendship might not be threatened by one party being incapable of keeping a secret. An open love affair might involve no commitment to excluding sexual intimacy with everyone else. But, realistically, many friendships, marriages, and loving partnerships, are essentially loyal relationships in the sense that they do require excluding others in some respect.⁶

The bank robbers in a gang are collaborating in a heist and share a commitment to the successful robbery and to getting away with it. This commitment requires further commitments – in particular, the commitment to not sharing the plot with anyone outside the gang. Sharing the plot would be working for the authorities against the gang and would be a breach of loyalty. In general, a group isn't a gang without a shared commitment to loyalty. A trainspotting club may have a shared goal and commitment to collaborate on it. But that goal won't require taking sides, and loyalty is not an essential aspect of being a member of that sort of club. A football club, on the other hand, being a competitive group, does require loyalty. If a player is persuaded by a gambling syndicate to throw a game they have breached that loyalty.

5. Conclusion

When you are accepted into a relationship or group that involves taking sides – an us/them relationship – the other members of the relationship or group are taking a risk. They are relying on you to do what that membership

⁶Whether this is a good thing or not is not my concern here. Keller (2007) argues that loyalty is not itself a virtue as relationships that depend on it, including patriotic allegiances, are not always particularly moral.

commits you to do. The success of the collaborative mission to rob a bank, for example, depends on everyone in the gang doing their part. In virtue of this being an us/them relationship, one aspect of your doing your part is your staying on side. Success of the mission depends on your loyalty. So, in accepting you into the relationship others are relying on your loyalty, making themselves vulnerable to your breaching that loyalty; they are trusting you to remain loyal. I have argued that betrayal is exactly a breach of that trust.

Suppose that Nick always cuts my hair. But one day, I need a haircut in a hurry and I go to his rival, Simon. I'm sure it would be irrational to feel anxious that I have somehow committed an act of betrayal; but I might nevertheless feel this. I would feel it because I had started to feel that the relationship I had developed with Nick demanded loyalty – that he was the only one allowed to cut my hair. And it would be irrational to feel this, because, in fact the relationship, while a friendly one, was not essentially an exclusive one. I had been a loyal customer, and perhaps Nick had come to rely on my loyalty. But it would be wrong to say that he trusted me to remain loyal, since our relationship did not involve any commitment on my part to remain loyal. And without this commitment, my 'disloyal' act was no breach of trust, and consequently not a betrayal.

Monogamous sexual relationships involve a system of entitlements and commitments, including a commitment not to entitle others to having a sexual relationship with you. In entering into such a relationship your partner trusts you to do your part in maintaining this relationship. You might breach that trust in all sorts of ways by failing to do what you are supposed to do given this system of entitlements and commitments. But not all such breaches of trust are betrayals. For example, your commitment to continue to love your partner may be breached without this being a betrayal. But your commitment to remain loyal – to retain the exclusivity of the relationship – has a special status. As this is a commitment that is essential to the relationship being what it is, your partner must be trusting you to remain true to it in accepting you into the monogamous relationship with them. A breach of this trust is what constitutes betrayal.

It is worth introducing the topic of jealousy briefly at this stage. Sexual jealousy seems to be the emotion that is the natural response to actual or threatened sexual betrayal. It is usually taken to be pathological in some way – an inappropriate response reflecting an excessively possessive attitude. But if there is some value in the relationship being monogamous and the betrayal or threat of betrayal is real, then some emotional response to that is surely appropriate.

The notion of jealousy tends to be limited to intimate relationships. But we can extend it to other betrayals. According to the Bible in the book of Exodus, God describes himself as a jealous god. This is linked to the demand that his followers shall have no gods but him. He is jealously guarding the

loyalty of his followers. And the emotion of jealousy is appropriate because he is in a compact with them, trusting them to remain loyal as their part in the collaboration. While we wouldn't describe the emotional response of gang members to the actual or threatened betrayal by another gang member as jealousy, it has a lot in common with sexual jealousy and with the jealousy of the Old Testament God.

Acknowledgments

I would also like to thank the members of David Charles' Discussion Group, my colleagues at UCD and especially the editors of this special issue for very helpful comments and discussions.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

Work on this publication has been funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No [870883].

References

- Baier, A. 1986. "Trust and Antitrust." *Ethics* 96 (2): 231–260. doi:10.1086/292745.
- Baker, J. 1987. "Trust and Rationality." *Pacific Philosophical Quarterly* 68 (1): 1–13. doi:10.1111/j.1468-0114.1987.tb00280.x.
- Cogley, Z. 2012. "Trust and the Trickster Problem." *Analytic Philosophy* 53 (1): 30–47. doi:10.1111/j.2153-960X.2012.00546.x.
- D'Cruz, J. 2018. "Trust within Limits." *International Journal of Philosophical Studies* 26 (2): 240–250. doi:10.1080/09672559.2018.1450080.
- Domenicucci, J., and R. Holton. 2017. "Trust as a two-place Relation." In *The Philosophy of Trust*, edited by P. Faulkner and T. Simpson, 149–160. Oxford: Oxford University Press.
- Faulkner, P. 2015. "The Attitude of Trust Is Basic." *Analysis* 75 (3): 424–429. doi:10.1093/analys/anv037.
- Fletcher, G. 1993. *Loyalty: An Essay on the Morality of Relationships*. Oxford: Oxford University Press.
- Holton, R. 1994. "Deciding to Trust, Coming to Believe." *Australasian Journal of Philosophy* 72 (1): 63–76. doi:10.1080/00048409412345881.
- Keller, S. 2007. *The Limits of Loyalty*. Cambridge: Cambridge University Press.
- Kleinig, J. 2022. "Loyalty." In *The Stanford Encyclopedia of Philosophy*, (Summer 2022 Edition), edited by E. N. Zalta, <https://plato.stanford.edu/archives/sum2022/entries/loyalty/>
- McGeer, V. 2008. "Trust, Hope and Empowerment." *Australasian Journal of Philosophy* 86 (2): 237–254. doi:10.1080/00048400801886413.

- McLeod, C. 2021. "Trust". In *The Stanford Encyclopedia of Philosophy*, (Fall 2021 Edition) . edited by E. N. Zalta, <https://plato.stanford.edu/archives/fall2021/entries/trust/>
- Noonan, H. 2021. "No Trust Is Hybrid: Reply to Faulkner." *Philosophia* 49 (5): 2189–2195. doi:10.1007/s11406-021-00341-7.
- Simpson, T. 2012. "What Is Trust?" *Pacific Philosophical Quarterly* 93 (4): 550–569. doi:10.1111/j.1468-0114.2012.01438.x.
- Thompson, C. 2017. "Trust without Reliance." *Ethical Theory and Moral Practice* 20 (3): 643–655. doi:10.1007/s10677-017-9812-3.