

TRUSTING ON ANOTHER'S SAY-SO

GRACE PATERSON

University of Vienna

We frequently trust others—even strangers—based on little more than the good word of a third party. The purpose of this paper is to explain how such trust is possible by way of certain speech acts. I argue that the speech act of vouching is the primary mechanism at work in many of these cases and provide an account of vouching in comparison to the speech act of guaranteeing. On this account, guaranteeing and vouching both commit the speaker to certain actions conditional on undesirable events. However, while guaranteeing may generate reliance in someone or something by providing insurance against bad outcomes, vouching may generate trust in another agent by way of the vouching party manifesting their own trust in the party vouched for and enjoining the addressee to trust likewise.

1. Introduction

Consider a rather ordinary scenario: you are going on vacation and need someone to take care of your cat. You ask a friend and they tell you that they cannot do it themselves because they too will be away, but that their good friend Sam is very good with animals. You could continue to cast around for help but, after some discussion, you instead go ahead and ask Sam to pet-sit for you. You do this in spite of the fact that Sam is a stranger about whom you know nothing except for what little your friend has told you.

Consider another scenario: you are out late and want to go home. Although you have alternatives available to you (you could stay the night on a friend's couch, for instance) you decide to hail a cab. You have never met the driver in your life, but they have a clearly marked vehicle and a taxi license on display. Seeing this, you let the cab take you home.

In this paper I shall argue that these are both examples of trusting based on the good word of a third party. In the first case, you trust Sam to care for your pet

Contact: Grace Paterson <grace.paterson@gmail.com>

based on what your mutual friend has said. In the second case, you trust the cab driver to treat you with dignity and take you to your home safely because of her relation to the company for which she works. Indeed, that relationship is likely the only thing you know about the driver. The cases have same general structure: at the outset, you trust someone or something (your friend, a company) but you do not yet trust someone else (your friend's friend, the cabby). In the course of events, the party you already trust says or does *something* that facilitates your coming to trust this other individual. That interaction is the focus of this paper.

I will argue that one of the most common and effective ways to generate such trust is by *vouching*. The thought is that, in some way—even if only tacitly—your friend vouches for Sam in the capacity of carer. Similarly, the cabby's taxi license and affiliation with a legitimate company serve, plausibly, to represent a vouch issued by a trusted authority. In both cases, then, the fact that a trusted party vouches for a stranger leads to trust in that stranger. In both cases, moreover, the trusting act is neither forced nor coerced: there are viable alternatives available which, although slightly less convenient, are safer in the sense that they do not involve stranger trust.

Vouching, then, is a speech act designed to generate trust in the agent vouched for. Moreover, it does this directly: that is, the vouch is not merely aimed at establishing conditions under which trust might blossom over repeated or sustained interactions. Rather, when successful, it allows the addressee to trust the one vouched for on the basis of the vouch alone. The point is vital because many instances of vouching apply to circumstances where there will be one and only one potentially high stakes interaction between the trusting party and the vouched for party.

In what follows I provide an account of vouching in comparison to a similar speech act: *guaranteeing*. I begin by distinguishing in §2 between reliance and trust, which represent the perlocutionary aims of guaranteeing and vouching respectively. I examine guaranteeing in §3 and build upon it to provide an account of full-on vouching in §4. The hope is that the analysis will illuminate how trust in strangers based on the word of a third party is possible.

2. Laying the Groundwork

Before we can understand vouching, we must understand what it is aiming for and why this aim is both important and difficult to achieve. I take it that the function of vouching is to enable us to trust those who we are not already prepared to trust. It does this by harnessing the trust we already have in the one who does the vouching in order to facilitate trust in the one vouched for. In this way, the speech act of vouching helps with the socially facilitated spread of trust

in a manner akin to how the speech act of testimony helps with the spread of knowledge. Like testimony, vouching is a tool of the social world that allows us to achieve an end even when we lack the resources to do so as an individual. Just as testimony helps us to acquire knowledge of things to which we are unacquainted or lack expertise, vouching helps us to relate to out-and-out strangers in a manner that we normally reserve for those closest to us. Just as testimony has its epistemic character because of its epistemic aim, vouching has a moral character derived from its moral aim. And just as an account of testimony must be sensitive to the requirements of knowledge, an account of vouching must be sensitive to the requirements of trust.

Thus, we will need to understand a little about trust and, in particular, what distinguishes it from the less taxing attitude of reliance.¹ Since there is no consensus on how to define either trust or reliance, it should be noted that the aim of this paper is neither to adjudicate the ongoing debates on this topic nor to provide a novel account of either of these attitudes. Nonetheless, we will need to have some degree of background theory in place in order to properly examine the speech acts of interest. In this section, I therefore will lay out what I take to be the most important features of these attitudes for the purposes of this project.

I will follow much of the literature in describing both trust and reliance as three place relations.² One trusts another *with* something or trusts them *within* or *with respect* to a particular domain.³ Similarly, I take it one relies on someone or something to perform some particular task, or for some particular function. Thus, I rely on the bus to be on time, the rope to hold my weight, or the waiter to come take my order. Sometimes, for the sake of perspicuity, I will leave the third argument in these relations undefined. Nonetheless, a domain specification is assumed in all cases. In our opening examples, you trust Sam with respect to caring for your pet and the taxi driver with the transport of your person. The trust in each case pertains specifically to the domain in question. You don't trust the cabby with your cat, for instance, and who knows if Sam can even drive.

1. A prevalent view is that trust involves, requires, or is partly constituted by reliance. This has not, however, gone uncontested—for instance, it has been argued that the attitudes can come apart (Thompson 2017) and even that trust can ground reliance rather than the other way round (Marušić 2017). I will remain neutral on whether reliance is, in some sense, a necessary or a part of trust. It is enough for our purposes that reliance and trust are compatible.

2. Domenicucci and Holton (2017) and Faulkner (2015) model trust instead as a two place relation: one simply trusts another person and without restriction. I am skeptical about this kind of unqualified trust, except perhaps in special cases such as the trust children have in their parents. Certainly it seems overly demanding for the purposes at hand. Even if such trust is possible, I think it unlikely that it is what is at work in most vouching. Those rare cases where it is applicable may be captured on the three place model by taking the domain of trust to be unrestricted.

3. I am not concerned here with the specific formulation of three place trust. One could follow authors like Baier (1986) by speaking of trusting someone *with* something, or those Jones (1996; 2012) and speak of trust *within* a domain. Either of these will work for my purposes.

Of course, this doesn't preclude the possibility that you might trust someone with something in another domain as well. The point is that trusting someone with respect to one domain does not entail anything in particular about other unrelated domains. Notice as well that while we must trust the individual doing the vouching, we typically trust in them in a different way from how we come to trust whomever is vouched for. Trusting a friend to vouch for someone requires that you trust them as a voucher, but it does not require that you trust them with respect to the domain the vouched for person is to be trusted in. You might not trust your friend to cat sit since she is terrible with animals herself. Nonetheless, you might still trust Sam for whom she has vouched in this regard. Similarly, you don't trust the Taxi company to itself take the wheel and drive you around (what would that mean?) although you trust its drivers to do so.

An important difference between trust and reliance, as I understand them, is that one trusts only other people (or, at least, other agents), but one can rely not merely on people but also on objects or events. So, for instance, I can trust my friend to hold my rope while climbing, but I can merely rely on the rope to be strong. Similarly, I can trust a neighbour to water the garden, but can merely rely on it to rain as forecast. Not unrelated to this is the observation that there are differences in terms of the significance of trusting or relying on an individual who fails to live up to expectation. As Baier colorfully puts it, "Kant's neighbors who counted on his regular habits as a clock for their own less automatically regular ones might be disappointed with him if he slept in one day, but not let down by him, let alone had their trust betrayed" (1986: 235). That is, trust can be betrayed whereas reliance can only be disappointed. I will gloss this facet of trust in terms of how the trusting party relates to the trusted party. I take it, following Richard Holton, that trusting involves adopting the participant stance toward the individual being trusted.⁴ We look upon a trusted individual as a fellow agent to be reasoned *with* rather than an object or obstacle to be reasoned *about*. This stance comes with a readiness to respond to the trusted agent's actions with distinctive reactive attitudes—attitudes which include that sting of betrayal unique to trusting situations.

Notice that merely because someone is an agent does not mean that every kind of dependence on them is a matter of trust. The fact that it would strike us as strange if Kant's observers were to describe his sleeping in as a betrayal suggests that they are not relating to him in the right way for their dependence to be

4. Holton (1994) more specifically argues that trust is reliance from the participant stance, although I do not require this tight a tie to reliance (see fn. 1). Holton is, of course, drawing on Strawson's "Freedom and Resentment" (Strawson 2008). For discussion see also Marušić (2017), Helm (2014) and citations therein. Related ideas are found already in Baier, who suggests that "one can only trust things that have wills, since only things with wills can have goodwills" (1986: 14), and Jones (1996) who holds that trust involves a kind of "optimism" about the trusted individual.

considered trust. They may set their watch by his daily promenade but they are not, for those purposes, treating his agency as a relevant factor. Instead, they are assuming that with respect to this matter he will continue to do precisely what he has done every other day. They are in essence treating him as if he is clock-work. It is for this reason that we are inclined to say that they are relying on him rather than trusting him.

Trusting often involves the truster granting the trustee a certain kind of discretionary power within the context of the trusting interaction. She grants the trustee the freedom to exercise her agency in ways that affect her, and acts on the assumption that the trustee will help rather than harm her. There is, in virtue of this freedom, an open-endedness to how we trust that contrasts with the more rigidly circumscribed form our reliance takes. This enables trust in situations characterized by at times a great deal of uncertainty and risk. When I trust my friend to take care of my elderly cat while I am away, I do so knowing that my cat could take ill while I am gone and the assumption that my friend will respond to this turn of events appropriately. Indeed, because it does not involve forming explicit expectations about what the trustee will do, trust is powerful—and sometimes indispensable—in contexts of unpredictability or when there is a salient expertise gap between truster and trustee (as there is when we visit a doctor).

It should come as no surprise, given these points, that trust tends to arise among individuals who are already acquainted. Generally speaking, the better we know someone, and the better they know us, the easier it is to depend on them to make decisions that are in our best interests (or not, if we come to distrust them instead). Ongoing relationships with repeated interactions therefore provide the conditions most conducive to trust. Strangers, by contrast, are comparatively difficult to trust even if we are able to rely on them in certain ways. We have no prior experiences with them and there is no particular reason to assume they will do especially well by us. In general if we are to trust strangers, we will require some help.⁵ That is the puzzle which animates this paper. My central claim is that vouching for someone is a way to enable trust in situations there is not sufficient time for it to develop organically through a relationship.⁶

To summarize, I take both trust and reliance to be attitudes tied to our practical reasoning. Relying on something or someone involves a willingness to

5. An exception of sorts exists for minimal forms of “civic trust” (Preston-Roedder 2017) examples of which include trusting other strangers “not to steal our belongings or push us onto the tracks” (2017: 4) when on the subway, and being unafraid to “ride a city bus with them or stand in line with them in a crowded market” (2017: 5). The cases that we are concerned with in this study occur in contexts of deeper interpersonal engagement where the trustee is expected to actively do things for the trusting party, rather than simply refrain from harming them.

6. It is, of course, not the only social mechanism that aims at fast tracking trust. Trust building exercises also aim at engendering trust among strangers or near strangers.

act based in part on certain assumptions about their behaviour or functioning. Trusting similarly involves a willingness to act based on certain assumptions, however these assumptions are related more deeply to facts about the trustee's character and agency and are made from the participant stance. Trust is, in this way, a substantially interpersonal attitude.

With these points in mind, we turn now toward the speech acts that interest us. The primary social function of these speech acts is to help a potentially hesitant addressee get to the point of relying on or trusting in someone or something. We begin with the guarantee which aims at facilitating reliance.

3. Guaranteeing

Guarantees can take many objects: one may guarantee a person or a machine, that a future event will obtain, or that a certain statement is true. Thus, for instance, a manufacturer may guarantee a toaster against damage through normal use. They may also guarantee that their toaster has been through a proper inspection process.

Guarantees are related to assertoric speech acts in the following way: in guaranteeing a product, the manufacturer seems to be making a descriptive claim about the product's reliability. Similarly, in guaranteeing that it has been inspected, they seem to be claiming that it has been indeed inspected.

But while guarantees certainly involve a descriptive component, they are not purely assertoric. We can see this by noting how they are tied to our practical reasoning in a manner that differs from pure assertions. As Turri puts it, "one main motivation for making guarantees is to provide others with enough assurance that they're willing to proceed with a course of action, in contexts where they aren't satisfied with mere assertion" (2013: 603).⁷ On my view, guarantees achieve this perlocutionary effect by including a commissive element. That is, guarantees are hybrid speech acts having both an assertoric and a commissive parts.⁸ The speaker aims at creating commitments for herself that are of

7. Turri is concerned with the connection between "I know that Q" statements (assurances) and guarantees. On his view, "by saying 'I know that Q' you represent yourself as having the authority to guarantee that Q" (2013: 607), where this requires satisfying a more stringent epistemic norm than with ordinary assertion. Specifically, "you may guarantee Q only if you know that you know Q" (2013: 608). I disagree with the applicability of this norm for guaranteeing in general, although it is plausible that it fits a particular strong kind of guarantee implicit in assurances. On this see also Lawlor (2013), especially her discussion of Austin's suggestion that assurances involve "unlimited" guarantees.

8. The idea that guarantees are hybrid speech acts is already found in Bach and Harnish (1979). Similarly, Turri (2013) suggests that guarantees might be classified as "assertive commissives". Note that the commitment in (ii) is practical: it is a (conditional) commitment to take certain

practical significance to the hearer. The guarantee of a manufacturer of their toaster against defects doesn't have to entirely convince you that the toaster has no defects, only that any worries about its being defective shouldn't prevent you from buying it. The manufacturer is, after all, committed to compensate you if it should prove defective.

I define guaranteeing as follows:⁹

In uttering *u*, *S* guarantees *X* in context *C* to addressee *A*, just in case *S*:

- i* asserts that *X* may be relied upon in context *C*, and
- ii* commits to appropriately compensate *A* if *X* fails in context *C*.

Another way to phrase (i) is to say that *S* asserts that *X* is reliable in context *C*. I have chosen this phrasing because it can accommodate all the various potential objects of guarantees. As part of guaranteeing the cabby, the taxi company asserts she may be relied upon to drive safely. As part of guaranteeing a toaster, the manufacturer asserts that it is reliable in standard toasting contexts. As part of guaranteeing that the toaster has been inspected, they assert that one can rely on the truth of the proposition "that the toaster has been inspected".

The role of context in this definition is to limit the scope and applicability of the guarantee so that it is not unrealistically demanding. Someone may be guaranteed to water my garden but not to fix my appliances, a toaster may be guaranteed when used to make toast but not as a door stop, and a weather forecast may be guaranteed to be accurate here but not there.

By appropriate compensation it is meant that the guarantor will compensate the addressee in a manner proportional to the harms or losses she sustains as a result of the failure of the guarantee.¹⁰ So if the addressee suffers greatly, the

actions. Depending on one's view of assertion, (i) may also carry with it certain commitments; however, these will be doxastic in character. One may, for instance, be committed to defending the claim that *X* may be relied upon in *C* by providing evidence about her reliability. On the difference between doxastic and practical commitments, see Brandom (1994: 238–43). Compare with: telic vs atelic commitments (Geurts 2019); direction of fit (Searle 1975).

9. I am basing this formulation roughly on Bach and Harnish who give the following schema: "Guarantee that: *S* affirms (constative) the quality of something, *x*, and promises to make repairs or restitution if *x* is relevantly defective" (1979: 50). Compare also: "when I guarantee I guarantee that, upon some adverse and more or less to be expected circumstance arising, I will take *some more or less definitive action to nullify it*" (Austin 1961: 70, emphasis in original); "guaranteeing that *Q* is true seemingly commits you to undertaking responsibility for other people's actions based on your words" (Turri 2013: 610).

10. An account of guaranteeing from a non-speech act perspective found in Coleman (1990: chs. 5, 8) contains similar ideas. Coleman does not distinguish reliance from trust, nor guaranteeing from vouching; however he does hold that what makes a guarantee distinctive is that the guarantor "experiences a loss of resources if the final trustee violates trust" (1990: 181) and that this laying down of resources plays an essential role in facilitating the transfer of (what he considers) trust. See also Hardin (2002: ch. 6) for related discussion.

guarantor will owe her more than if she is only mildly inconvenienced. Notice as well that in some cases, the guarantor may also commit to some minimum level of compensation to the addressee in the event that the addressee's losses cannot reasonably be compensated in full. As we shall see, however, such a guarantee is of more limited effectiveness and will sometimes be insufficient to motivate the addressee as intended.

Now certain speech acts, such as commands, baptisms, and legal speech acts, sometimes also require that the speaker have a particular entitlement or authority if they are to be successful. We should consider briefly whether a guarantee is such a speech act. One reason to suspect it is is that there seems to be something strange about certain parties issuing certain guarantees. For instance, it seems peculiar for a company to guarantee some product made by an entirely different company. This would suggest that perhaps only certain individuals, such as the maker of a particular product, are entitled to guarantee it.

I do not believe that is correct. The peculiarity here does seem to me to derive not from a lack of entitlement. Rather the choice to issue a guarantee is somewhat suspect such that we might reasonably worry that the speech act will be infelicitous in other ways. The speaker is, after all, less likely to have the knowledge required to make the judgement communicated through a guarantee. Similarly they may be much less motivated to follow through on the guarantee than those who are more intimately connected with the success of the object of guarantee. Nonetheless, there are still many cases in which guarantees could be made in absence of any special connection between the guarantor and what they are guaranteeing. One could, for instance, issue a guarantee to an anxious friend that the rope she is dangling from is very sturdy or that her book bag is water resistant enough to withstand the rain. Such guarantees can indeed have a reassuring effect. The point is that a guarantee generally requires no more special authority than any other commissive speech act,¹¹ although there may also be some institutionalized variants of guarantees which may be externally enforceable.

Condition (ii) encodes the conditional face of a guarantee. This is worthy of further examination, because there is something that makes the commissive aspect of guaranteeing (and, we shall see, vouching) quite unusual: they are *conditional commissives*. Standard commissive speech acts, epitomized by the act of promising, involve the speaker committing to some course of action. For this to be felicitous, the speaker should intend to perform the action in question. Thus if Sam promises to care for your cat while you vacation, she should intend to care

11. That is, they require no stronger an entitlement than is required for a commissive such as promising. To borrow the vocabulary of Kukla and Lance (2009), these speech acts do not have entirely agent-neutral inputs; but that is just because the speaker is the only one entitled to generate such commitments for themselves.

for your cat and plan accordingly. If she does not so intend, there is something wrong, indeed insincere, about the promise she has made. This is not the only condition upon the felicity of such a speech act, but it is an important one.

Conditional commissive speech acts, by contrast, involve a commitment to take an action *if* something specific occurs. Moreover, in the case of guarantees, this “something” is undesirable. These facts change what is expected of the speaker. Specifically, she is expected only to have a conditional intention: something of the form “I intend that if X happens, I will A”. She should plan to act in a way that is compatible with taking the action should it be called for, and she should be attentive to whether or not the condition is triggered. But it is entirely possible that, in fact, no further action will be required of her.¹²

As a consequence, these conditional commissive speech acts have certain further felicity conditions: the addressee should not desire the condition obtain, and the speaker should neither believe nor desire for it to obtain. Indeed, to have such a desire would be to root for something that is against the interests of the addressee. At the same time, the commitment itself should be in the interests of the addressee since it is supposed to be compensation for their bad fortune. As Searle noted, “a promise is a pledge to do something for you, not to you; but a threat is a pledge to do something to you, not for you” (1969: 58). In just the same way, a conditional pledge to do something *to* rather than *for* you would constitute an ultimatum, not a guarantee.¹³

So it is an important part of a felicitous guarantee that the guarantor neither expects nor desires their commitment to ever have to be discharged. They intend that if things go wrong, they will take the promised actions to remedy this situation, but they also intend (to the degree that they can) that this circumstance not arise. Since a guarantee made with the belief or expectation that it will actually be necessary to take the promised actions is to some degree infelicitous, an ideal guarantor is absolutely committed to taking remedial steps in the event they have to, but also absolutely certain that they will not have to. For this reason, not only failing to act on your guarantees but also having to act on your guarantees too often can undermine your own reputation as a guarantor. A guarantor having to follow through on their commitment is an indication that something has gone amiss.

12. Excellent analyses of conditional intentions can be found in Ferrero (2009) and Ludwig (2015). I have also benefited from discussion with Matthew Rachar who considers them in Rachar (2021).

13. One way to understand the speech act of threatening is that the speaker commits to doing something the hearer does not want on the condition that the hearer does something the speaker does not want. As with a guarantee, the speaker’s preferred and intended outcome involves not having to follow through on her commitment, but she should all the same be prepared to do so. For details see Schiller (2019).

The upshot is that the guarantor should not intend to perform the action they have committed to. If anything, they should actually intend to not perform the action promised while at the same time being willing and prepared to do so if they must. Given, however, that the commitment is not intended to be acted upon, it is not clear how we should understand its role. Put differently, the puzzle is to explain how a conditional commissive of this sort can make a practical or rational difference given that it is expected to lie essentially dormant. Let us turn now to this problem.

Recall that the guarantee aims at generating reliance. The idea here is that the presence of the conditional commitment on the part of the guarantor alters the calculus of risk related to the object guaranteed in a way that may be able to tip the balance for the addressee and allow her to rely on it. The commitment made does this in two ways: first, it gives reason for the addressee to lower her expectations that the failure state will obtain. This is because the guarantor has signalled confidence in what they have guaranteed. Given that there is potential cost to them, the thought goes, it would be not only infelicitous but indeed irrational for the guarantor to issue the guarantee if they expect things were likely to go awry. So the very presence of a guarantee is at least some evidence that the object of the guarantee is reliable.

The second way in which the guarantee alters the calculus of risk is that it reduces the cost of failure should it occur. This is, of course, because the guarantor has committed to compensating losses. Consequently, in the off chance things do go wrong, it won't be especially bad for the addressee—or at least it won't be as bad as it could have been were the guarantee not in place. Notice that this aspect of a guarantee allows for the possibility that the addressee could be, to some degree, unconvinced by the assertoric face of the guarantee. She might have some lingering doubts about the quality of the toaster she buys but go ahead with the purchase because the guarantee gives her a back-up plan.

Put differently, a speaker issuing a good faith guarantee should have confidence in reliability of what they guarantee; but the addressee may nonetheless harbour doubts. The commitment attending the guarantee helps in these cases. It means that even if the addressee is not fully convinced that what has been guaranteed is as reliable as the guarantor claims, they may still have good reason to act on the assumption that it is. One risk that the commitment helps defray, then, is the possibility that the guarantor could be wrong in the assessment they make.

So one has reason to believe that relying on that which was guaranteed is both less risky and, in the worst case, less costly. The guarantee acts as a kind of insurance. It is in this way that a guarantee can help induce reliance on a stranger. A guarantee that someone will move the addressee's belongings safely may move the addressee to rely on them to do so, a guarantee of someone's ability to care for plants may move the addressee to rely on them to plant sit, and a

guarantee of a doctor's competence may be enough for a patient to consult them about, say, a minor wound.

However, the guarantee suffers from some inadequacies which limit its applications in our social lives. One issue is that a guarantee may not alter the calculus of risk *enough* to move the addressee to action with respect to what has been guaranteed. Indeed, it may not be possible for it to do so. There are practical limits on how much a guarantee can reduce the potential costs of a particular act of reliance. As such, a guarantee may suffice to convince an addressee to let a stranger move their IKEA furniture, but probably not their irreplaceable heirlooms. It may convince them to let a stranger take care of their plants, but probably not their beloved cat. And it may be enough that they let a new doctor dress their wound, but not perform open heart surgery. Indeed, when a guarantee's failure would lead to the addressee's own death then a commitment to compensate the addressee can hardly be said to offer a satisfactory back-up plan. It is quite hard to adequately compensate a corpse.

There are various ways, of course, to try to adapt guarantees to bolster their effectiveness in these kinds of circumstances. One that is sometimes seen is to direct the promise of compensation toward other interested parties—perhaps the addressee's family. Such adjustments can lend credibility and seriousness to a guarantee that might otherwise fall flat. By promising to compensate others should something befall the addressee, the guarantor ensures she cannot trivially escape her responsibilities by having the addressee perish.

But these strategies can do only so much. Ultimately, in these cases, only one of the two deliberative fronts on which guarantees operate remains deliberately salient to the addressee. While the guarantor's assurances that the risks are fairly low may be credible, the potential for significant and unmitigated loss lingers.

Another limitation guarantees face is more fundamental. Given what we have said about trust involving engagement with an agent *as* an agent, there is a real worry that a guarantee works at cross-purposes to trust. Put plainly, it seems wrong to say that you trust someone if, in fact, you lean on someone else entirely to mitigate the risks associated with such dependence. The way that the addressee leans on a guarantor is downright antithetical to actually trusting in the person guaranteed.

This is, I think, at least in part because the purely practical considerations supplied by a guarantee are the wrong kinds of reasons for trust (Hieronimi 2008). A guarantee encourages the addressee to engage with the object of the guarantee on a purely instrumental level: she evaluates the likelihood that that person will do this or that, takes into account the potential wins and losses, and calculates her best possible response as a result. She does not, as part of this process, need to engage with the guaranteed agent as an agent. How she reasons

when considering a guaranteed person is not fundamentally different from how she reasons when considering a guaranteed toaster.

Vouching, to which we turn now, attempts to generate trust rather than reliance. It does this in part by offering reasons that are not purely instrumental in nature. Although it is structurally similar to a guarantee, it encourages the addressee to engage interpersonally with the vouchee herself and to do so without relying on the voucher to supply a safety-net.

4. Vouching

Like a guarantee, a vouch is a hybrid speech act with both assertoric and conditionally commissive elements. In vouching, the speaker manifests her own trust and confidence in the vouchee to do well by the addressee. This has the effect of essentially enjoining the addressee to also trust the vouchee. How the voucher manifests her trust is that she voluntarily subjects herself to sufficiently severe potential loss in the event that the addressee is let down by the vouchee. This is not the same as committing to compensating the addressee—in fact, there is no commitment to directly changing things for the addressee at all. Instead, the speaker voluntarily takes on risk to herself thus binding her fate to that of the addressee. The idea is that even if it is not possible for the voucher to make up for potential harms the addressee might incur, it is nonetheless in her power to make herself similarly, although perhaps not equivalently, vulnerable. This is not an insurance policy in the way the guarantee is. Instead, the driving metaphor behind a vouch is that the voucher holds the addressee's hands.

I define vouching as follows:

In uttering *u*, *S* vouches for *X* in context *C* to addressee *A*, just in case *S*:

- i* asserts, based on second-personal knowledge of *X*, that *X* may be trusted in *C*.
- ii* commits to suffer a penalty commensurate with *A*'s losses if *X* fails in *C*.

A vouch resembles a guarantee insofar as it is a complex speech act with conditionally commissive element. It diverges from a guarantee in several important respects. First, as one would expect, its assertoric element pertains to the trustworthiness of *X* rather than merely the reliability of *X*. Importantly, and less obviously, the voucher must base this judgement on their own second-personal knowledge of the vouchee. By this I mean that she must herself *know* the vouchee in the distinctive sense that one person knows another. It is, that is, not

enough for her to know various facts *about* the vouchee; she must have engaged with the vouchee as a subject, and these interactions must be at least part of why she has judged the vouchee to be trustworthy.¹⁴

Not only does this strengthen the credibility of the vouchee's claim in a way other kinds of third person evidence cannot, it also allows her to present the vouchee to the addressee from the participant stance—the stance appropriate for trusting. In this way, vouching is much like an act of introducing one person to another. When we make introductions between two otherwise strangers, we invite them to acquaint themselves with and ultimately become known to one another. Vouching can similarly help launch a trusting relationship between strangers.

An important consequence of this constraint is that one cannot felicitously vouch for a stranger.¹⁵ This I believe is a welcome result. Were that the case, then we would see chains of vouching wherein one party could vouch for another party based only on someone else's vouch. But I believe that this is intuitively odd. The reason for that oddness is that part of what one is offering up with a vouch is a kind of first-hand experience, presumably a kind of experience that the addressee herself lacks. Vouching therefore generally occurs in cases where there is a kind of epistemic gap between voucher and addressee: the voucher knows the vouchee better than the addressee does. Indeed, it would be very strange (although not entirely infelicitous) to vouch for someone that the addressee clearly knows quite well. One could try vouching for the addressee's mother, for instance, but I suspect it would be unlikely to achieve its perlocutionary aims. It is likely that either the addressee trusts her mother already, or that she has reasons for distrust that will be hard to overturn.¹⁶

Notice as well that condition (i) rules out cases of vouching for acquaintances where one's assertion of their trustworthiness is not based on your own second-personal knowledge of them. That is, it is not enough to know the individual you vouch for; your experiences with them must be what forms (at least in part) the grounds for your belief in their trustworthiness. I cannot felicitously vouch for someone I met recently solely on the basis of a report from a mutual friend. I must come to see their trustworthiness for myself, through my own interactions with them.

As with a guarantee, a vouch contains a conditional commitment which is meant to not obtain. In this case, the voucher commits not to providing compensation to the addressee in the case of failure but rather to suffering in a manner

14. On this kind of second-personal knowledge of others, see Benton (2017) and Talbert (2015).

15. I am grateful to an anonymous referee for drawing this datum to my attention.

16. On the other hand, I do not think that the speaker and addressee need to necessarily know each other for a vouch to be felicitous, although vouches by speakers unknown to the addressee are much less likely to succeed at their perlocutionary aim. This is because the addressee is more likely to lack the kind of confidence in the voucher that is required to take her at her word.

that at least compares to how the addressee herself suffers. This is important because, as we have noted, cases which call for a vouch often involve potential losses for which there is simply no adequate compensation. This is not, however, meant to be read as committing to a kind of self-punishment; it is more akin to the laying down of collateral. It is a demonstrated willingness to put one's own skin on the line.

There are similar sorts of non-speech act ways to commit yourself. For instance, one way to convince someone to trust your friend Sam to care for their cat is to leave your own cat with Sam as well. If I am unsure about letting someone transport my precious belongings, you could load yours in the truck alongside mine. These are acts of leading and trusting through example. So too is vouching. Vouching for someone is an expression of the speaker's own trust in them.

What the voucher puts on the line can take very different forms from case to case and needn't be of the same sort as what the addressee is considering putting on the line. What matters is just that it is something she clearly values. It may take the form of a potential financial burden, or it may take a form which is more intangible. When the voucher has themselves a tight relationship with the party that they are vouching for, that relationship may be what they put on the line. Thus the more intimate the existing relationship between voucher and vouchee, the stronger the force of the vouch even without having to invoke any other kind of collateral. In a different vein, an institution issuing a vouch will risk, among other things, their reputation. The import of this should not be underestimated: institutions that issue vouches often exist almost entirely for this purpose; to suffer significant reputation loss can therefore endanger their very existence.

Clearly, one difference between vouching and guaranteeing is that the object of a vouch must be an agent. Were it not, then the assertoric component of the speech act could not be satisfied. As a result, a vouch cannot be applied to an object or event the way a guarantee can. All of this, of course, echoes our earlier observations about reliance and trust. One may vouch for a friend, but not for, say, a product. A cocky forecaster can guarantee that tomorrow it will rain, and a manufacturer can guarantee their toaster against defects; but the forecaster cannot vouch for tomorrow's rain, nor can the manufacturer vouch for the toaster they have produced. Because of this, purporting to vouch for a non-agent will likely result in the addressee charitably interpreting it as a guarantee instead. A guarantee is, in many such cases, the closest speech act that could felicitously be issued.

On the other hand, it is not at all incoherent to guarantee rather than vouch for another agent. One might even vouch for an agent *and* issue a potentially measured guarantee about their behaviour (that is, a guarantee committing to compensation to the degree possible). Alternately, one might instead issue

a simple assertion about the character of another agent, without elevating this speech act to the level of either a guarantee or a vouch.

This raises the problem of how we are to distinguish between these speech acts in practice. Here I think the specific words chosen are not a reliable guideline. In ordinary usage the words “guarantee” and “vouch” are used fairly interchangeably, alongside other terms like “swear”. In practice, then, we look instead to how these speech acts are backed up to identify them. For instance an ad for toasters reading “quality toast or your money back!” is clearly a guarantee. We can see this in part because of its object (a non-agent) and the fact that it is offering up compensation in the case of burnt toast. A guarantee of a person may be backed up by a promise to help the addressee in the event of a problem caused by their not being reliable: “Kant won’t sleep in. If he does, you can come to me and borrow my own very precise watch.”

In contrast, “I swear on my life that this man may be trusted” is clearly a vouch. The speaker is offering not compensation, not a back-up plan, but collateral. Perhaps more subtly, suppose your friend supports her suggestion that you hire Sam to cat sit by talking to you about how well she knows Sam, how compassionately she has seen Sam behave toward animals, how she herself uses Sam’s services, and so forth. Here she seems to do more than just assert Sam may be trusted—but she does not offer a guarantee. We might be inclined to think she is, instead, vouching for Sam given that she is supplying a flurry of interpersonal reasons to believe Sam trustworthy. While the collateral isn’t made explicit, she might take pains to assure you that Sam betraying you would absolutely shatter her as well. In general then it is how the speech act is supported that tells us whether we are dealing with a guarantee, vouch, or simple assertion.

It is interesting to note that the reasons to trust that a vouch creates for the addressee can linger even in the absence of the voucher. This is not true of the reasons generated by a guarantee: the guarantor remains integral to the reliance that the guarantee produces. Should the guarantor vanish for some reason, the addressee’s disposition to rely would likely follow—unless, of course, other reasons to rely on whatever or whomever had been guaranteed were to come to the fore. By contrast, if the voucher vanishes, this will in no way undermine the fact that she was willing to take on the relevant risk. She has still demonstrated trust in the vouchee, and tied that trust to the fate of the addressee. Assuming that her absence is not due to something like an attempt to flee her own responsibilities, the fact that she is not there to see things through does not undermine the force of her vouch.

Similarly, the reasons created by vouching may have some applicability outside of the narrow scenario to which the vouch has been supplied. That is, one who has, in virtue of a vouch, come to trust another to do one thing may, in certain cases, also reasonably trust that person to perform other actions in virtue

of that same vouch despite those actions falling outside the domain of the vouch itself. This is because the vouch speaks to at least some degree to the character of the vouchee.

This is not true of the reliance generated by a guarantee, especially one directed at an event or object. The fact that my toaster has been guaranteed to function under normal use conditions gives me little reason to rely on it as a doorstop. Similarly, if someone is guaranteed to be a good plumber, I am not on that basis going to rely on them to take care of my cat. These things fall outside the scope of the guarantee issued and I have no reason to suppose that the person or thing is reliable in this new way, except to the degree that this overlaps with what they were guaranteed for. Moreover, there is nothing to mitigate the cost of the risks I may incur. A guarantee has little power beyond its narrowly defined scope of application. A vouch, because it may launch a relationship between two strangers, has much greater power.

With these points in mind, let us briefly revisit the limitations that guarantees faced and see how vouching differs on these fronts: they may not be effective in high stakes cases, and the presence of a guarantor might actually preclude or be incompatible with trust.

With regard to high stakes cases, the concern was that guarantees may fail because there are certain situations in which there simply is no compensation sufficient to make up for the potential loss the addressee is being asked to risk. Loss of one's own life is a particularly good example of this since in such cases the addressee would not even be able to claim whatever redress was on offer. Since guarantees function mainly by altering the calculus of risk for the addressee, they are largely useless when the risk is especially great or compensation untenable.

We have seen that vouching works on a different kind of logic, however. It involves the voucher taking on risk for themselves rather than offering potential compensation to the addressee. Because of this, whatever reasons it generates do not come from offsetting the dangers faced by the addressee. Instead, the addressee is encouraged and indeed invited to trust in part on the basis of seeing the voucher themselves trusting alongside them. This kind of reason is actually strengthened as the stakes increase; for the voucher will likewise take on greater risk themselves, making their own display of trust that much more powerful.

The second worry was that the structure of a guarantee is antithetical to generating proper trust. A guarantee works by creating a kind of safety net such that the addressee's existing reason to rely on the guarantor extends to cover her reliance on something or someone else as well. A guarantee, in other words, facilitates reliance on something that one would not have relied on, but only by tying the ongoing reliability of the thing guaranteed to a known quantity. A guaranteed agent is treated purely instrumentally: they are nothing more but a risk offset.

In the case of a vouch the addressee does not lean on the voucher in the same way as she would lean on a guarantor. Vouching specifically aims at getting the addressee to relate to the vouchee differently: to see her as an agent who will make decisions in her best interests. This is why it is conducive to trust. It creates an independent bond between the vouchee and the addressee. Once the participants have been so acquainted, the voucher herself is no longer needed.

Acknowledgements

A special thanks goes to Jorah Dannenberg for holding a seminar at Stanford that birthed an early version of this paper. Since then, the ideas have developed with generous help from many people on both sides of the Atlantic. I am particularly grateful to the participants at the 2019 workshop on “Group Agency and Collective Responsibility” at the University of Vienna, and to the reviewers and editors at *Ergo* for their hard work and constructive comments.

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 740922).

References

- Austin, J. L. (1961). Other Minds. In J. O. Ormson and G. J. Warnock (Eds.), *Philosophical Papers* (44–84). Oxford University Press.
- Bach, Kent and Robert Harnish (1979). *Linguistic Communication and Speech Acts*. MIT Press.
- Baier, Annette (1986). Trust and Antitrust. *Ethics*, 96(2), 231–60.
- Benton, Matthew A. (2017). Epistemology Personalized. *The Philosophical Quarterly*, 67(269), 813–34.
- Brandom, Robert (1994). *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard University Press.
- Coleman, James S. (1990). *Foundations of Social Theory*. Harvard University Press.
- Domenicucci, Jacopo and Richard Holton (2017). Trust as a Two-Place Relation. In Paul Faulkner and Thomas Simpson (Eds.), *The Philosophy of Trust* (149–60). Oxford University Press.
- Faulkner, Paul (2015). The Attitude of Trust Is Basic. *Analysis*, 75(3), 424–29.
- Ferrero, Luca (2009). Conditional Intentions. *Noûs*, 43(4), 700–741.
- Geurts, Bart (2019). Communication as Commitment Sharing: Speech Acts, Implicatures, Common Ground. *Theoretical Linguistics*, 45(1–2), 1–30.
- Hardin, Russell (2002). *Trust and Trustworthiness*. Russell Sage Foundation.
- Helm, Bennett W. (2014). Trust as a Reactive Attitude. *Oxford Studies in Agency and Responsibility*, 2, 187–215.
- Hieronymi, Pamela (2008). The Reasons of Trust. *Australasian Journal of Philosophy*, 86(2), 213–36.

- Holton, Richard (1994). Deciding to Trust, Coming to Believe. *Australasian Journal of Philosophy*, 72(1), 63–76.
- Jones, Karen (1996). Trust as an Affective Attitude. *Ethics*, 107(1), 4–25.
- Jones, Karen (2012). Trustworthiness. *Ethics*, 123(1), 61–85.
- Kukla, Rebecca and Mark Norris Lance (2009). 'Yo!'and'Lo!': *The Pragmatic Topography of the Space of Reasons*. Harvard University Press.
- Lawlor, Krista (2013). *Assurance: An Austinian View of Knowledge and Knowledge Claims*. Oxford University Press.
- Ludwig, Kirk (2015). What Are Conditional Intentions? *Method: Analytic Perspectives*, 4(6), 30–60.
- Marušić, Berislav (2017). Trust, Reliance and the Participant Stance. *Philosopher's Imprint*, 17(17), 1–10.
- Preston-Roedder, Ryan (2017). Civic Trust. *Philosopher's Imprint*, 17(4), 1–23.
- Rachar, Matthew (2021). Conditional Intentions and Shared Agency. Manuscript in preparation
- Schiller, Henry Ian (2019). Is That a Threat? *Erkenntnis*, 86, 1161–83.
- Searle, John R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.
- Searle, John R. (1975). A Taxonomy of Illocutionary Acts. In K. Gunderson (Ed.), *Language, Mind and Knowledge* (344–69). University of Minnesota Press.
- Strawson, Peter F. (2008). Freedom and Resentment. In *Freedom and Resentment and Other Essays* (1–28). Routledge.
- Talbert, Bonnie M. (2015). Knowing Other People: A Second-Person Framework. *Ratio*, 28(2), 190–206.
- Thompson, Christopher (2017). Trust without Reliance. *Ethical Theory and Moral Practice*, 20(3), 643–55.
- Turri, John (2013). Knowledge Guaranteed. *Noûs*, 47(3), 602–12.