

## 13

Trusting a Promise and  
Other Things*David Owens*

We are social animals and the form of our social life depends on features of our psychology. Some social theorists fix on a specific attitude as being crucial. For Hume the psychological cement of human society was sympathy; for Hobbes it was fear. For many recent writers, both popular and philosophical, trust plays this foundational role. According to the latter there is a special attitude we adopt towards other people when we trust them and that attitude is a *sine qua non* of our social institutions. So long as we trust one another, basic forms of cooperation will survive; where that attitude weakens or decays, things quickly fall apart.

I agree that society could not exist in its present form unless people both trusted one another and showed themselves worthy of that trust. Ubiquitous distrust would indeed be a social solvent (Reid 2010: 334–5). What I doubt is that there is a distinctive attitude of trust that serves as a psychological underpinning of our sociality; rather different forms of trust are appropriate to different sorts of object. All of these phenomena may be called forms of trust but to uncover their psychological character we must look not to the nature of trust (or distrust) but to the disparate character of the objects of trust.

## 1. A Thin Theory of Trust

Richard Holton maintains that much as we believe or intend that *p*, we may also trust that *p*:

Trust, I will suggest, is a distinctive kind of attitude involving a distinctive state of mind. My project is to look at the ways in which it is distinctive, and the ways in which it interacts with belief and with the will. (Holton 1994: 63)

For Holton trust, like belief and intention, is a propositional attitude with a specific psychological nature.<sup>1</sup> Given the ubiquity of trust, were Holton's assumption correct,

<sup>1</sup> 'Trust is an important and distinctive kind of psychological state' (McMyler 2011: 113).

we would expect the psychology of trust to be a fundamental element of our social theory and we'd also expect the value of trust to ground the authority of various social norms including those that require us to keep our promises.

Annette Baier shares these expectations and in a paper published some thirty years ago she expressed surprise at the 'strange silence' to be found on the topic of trust:

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... But when we turn to the great moral philosophers, in our tradition, what we find can scarcely be said to be even a sketch of a moral theory of trust. (Baier 1986: 232)<sup>2</sup>

More recent writers have sought to fill the gap, proposing accounts of the general nature and value of interpersonal trust. To complete their project these writers need not offer us an analysis of trust nor decompose the attitude of trust into other attitudes like belief and desire; trust may be indefinable, a psychological primitive. But if trust is a propositional attitude with a definite psychological character then it must be possible to answer some basic questions about it.

First, what kind of thing is trust? Is it in fact a propositional attitude or is it more like a speech act? If an attitude is it cognitive (like belief) or conative (like intention) or perhaps something in between like an emotion (Jones 1996)? I'll call this the *modal question*. Second, what is the characteristic content of trust? What must the trusted object be like for your trust in it not to be misplaced or mistaken? In the context of the recent philosophical literature on trust this question of content becomes what I'll call the *motivational question*. Third, what kinds of considerations justify trust? Is trust justified by evidence for the truth of some proposition or by considerations of a rather different sort or by some combination of the two? This is the *question of justification*. All of these questions have been raised in the literature on trust and none have been resolved to general satisfaction. Faced with this impasse, one may wonder whether 'the great moral philosophers' were not wise to refuse trust a foundational role in their ethical and social theories.

I maintain there is no such thing as the value of trust. Rather various objects of trust have their distinctive form of value and *for X to trust Y is for X to engage Y in a way that will realize the value distinctive of Y-things (provided this thing is a good Y)*. Call this a *thin theory of trust*. On this view, trust has no distinctive value; rather its role is simply to realize the value of the object of trust.<sup>3</sup> Furthermore, trust has no distinctive psychological nature for the attitudes that would realize the value of the various objects of trust are themselves very various. In some cases the only propositional attitude

<sup>2</sup> Most of what we do find is discussion of trust in a promise and, as Baier rightly says, that is just one form of trust.

<sup>3</sup> Contrast trust with pleasure. Even if the value of enjoyment is conditional on the value of the object enjoyed, there is a distinctive value in enjoying things: we should seek out (worthy) things to enjoy so that we have enough pleasure in our lives. We shouldn't in the same way seek out things to trust so that we have enough trust in our lives. The point of trust is simply to realize the value of the trusted thing, not to enhance that value by making us feel trusting towards it.

involved in *a* trusting Y may be an intention to behave towards Y in a certain way, though in other cases some further attitude (perhaps some motive for forming the relevant intention) might be needed to realize Y's distinctive value. And we can acknowledge all this without multiplying senses of 'trust'. For X to trust Y is always for X to engage Y in a way that satisfies the above specification.

Let's leave the realm of interpersonal trust for a moment. I trust my car when I am prepared to drive it around (not merely sit in it), an attitude that enables me to realize the distinctive value of a car. I trust an apple when I am prepared to eat it (rather than use it as a football), an attitude that enables me to realize the distinctive value of an apple. A specific apple and a specific car are worthy of this trust when they actually possess the value distinctive of their kind.<sup>4</sup> In the absence of trust, the value of a (trustworthy) apple and of a (trustworthy) car will be wasted. There is some leeway and a potential source of disagreement in the notion of a thing's 'distinctive value': am I trusting my apples if I'm only prepared to use them as compost? I shall offer no account of what might make edibility the distinctive value of an apple. I note only that disputes on this score are not to be settled by investigating the nature or value of the attitude of trust; rather it is the nature and value of apples that requires our attention.<sup>5</sup>

Trust is one of a range of psychological phenomena (including respect and appreciation<sup>6</sup>) the function of which is to realize the value of their object and whose nature and value is consequently as various as the nature and value of their objects. My main concern is to place trust in this class—it is less important to me how exactly we distinguish trust from its other members—but I'll say something tentative about what differentiates trust from respect, appreciation, and so forth.

One can realize the value of something in various ways without *engaging* that value in the sense I have in mind. For example, I might show respect for the value of the life of Joe the Hermit by deliberately avoiding any contact with Joe and I might show my appreciation of the value of a joke by laughing at it, or of a sunset by looking at it.<sup>7</sup> Here my attitudes of respect and appreciation help realize the value of their objects, a value which might otherwise be wasted but it would be strange to describe me as trusting the hermit, the sunset, or the joke. Trust implies a more positive engagement with its object, involving the expenditure of energy or one's undergoing a significant change that risks being misdirected unless the trust is well founded. Activity is required to realize the distinctive value of food, of machinery, of books, and so forth; even simply believing what someone tells you involves undergoing a complex psychological transformation with various emotional and behavioural ramifications. Boundaries are hard

<sup>4</sup> Trusting Y may or may not involve representing Y as trustworthy. That will also depend on whether such an attitude is required to realize the distinctive value of a trustworthy *object*.

<sup>5</sup> There may be objects that are valued by us but which have no *distinctive* value; they are just valued in different ways by different people. Perhaps the moon is such an object, valued by some as a celestial adornment and by others for its gravitational pull. In that case there would be no such thing as trusting the moon. We may trust *that* tonight will bring a full moon but we won't thereby be trusting the moon.

<sup>6</sup> Raz treats respect and appreciation as value-realizers (Raz 2001: 154–7).

<sup>7</sup> Thanks to Alison Hills for the hermit and Jessica Moss for the joke and the sunset.

to draw—if I strain to hear your jokes then at some point I am trusting your sense of humour, not merely appreciating it—but the point remains that trust is like respect and appreciation in that it realizes the value of its object.

How is this conception of trust to be applied to interpersonal trust? It is doubtful whether people as such have a distinctive value, one that can be realized by trust. Some might conclude from this that trust in things and trust in people are just two different phenomena. I instead propose that trust in a person be understood as shorthand for trust in something that a person is or does.<sup>8</sup> We think of various actions and character traits as having a distinctive value and this value is often wasted unless we respond to them with trust. You realize the value of my promises and my assertions by trusting them.<sup>9</sup> Character traits like good judgement, technical skills, and aspects of temperament (e.g. equanimity, self-confidence, self-discipline) are a similar case. You benefit from my driving skills or my courage (in the relevant way) by trusting them.<sup>10</sup>

Here I'll be focusing on the relation 'X trusts Y'. There are other idioms like 'X trusts that Y will A' which are often used to mean little more than 'X is relying on the fact that Y will do A'. For example trusting a promise is a different matter from trusting that a promise will be kept (perhaps because the police will force the promisor to keep it). One trusts *that* a promise will be kept whenever one extracts some value or other from the promise by relying on its being kept. To trust the promise (and thus the promisor) is, I shall propose, to engage with it in a way that realizes its distinctive value qua promise, a trust that (as we shall see) involves more (and perhaps also less) than simply relying on its being kept.<sup>11</sup>

I'll offer no general account of what makes V the distinctive value of Y but I shall say something about those objects of trust which (unlike apples) are governed by norms. Such an object's distinctive value explains the norms that govern it, so the distinctive value of a speech act (like a promise or an assertion) is the value which accounts for the characteristic content of the norms governing the speech act (e.g. those norms that distinguish promises from assertions and vice versa) and which explains why we should take those norms seriously.<sup>12</sup> For example, suppose the characteristic point of an assertion is to provide its audience with a way of knowing the proposition asserted. That will explain why assertions should be made on the basis of evidence for the

<sup>8</sup> So you want to be able to trust those around you because you want to be able to trust their assertions, their promises, their judgement, and so forth.

<sup>9</sup> You may be prepared to trust someone's promises but not their assertions; their honesty but not their courage. My account suggests why a refusal to trust people is often problematic.

<sup>10</sup> To maliciously expose me to danger is to exploit my courage rather than to trust it because it is to use my courage in a way that does not realize the distinctive value of courage. Similarly for exploiting my goodwill to trick me and the like.

<sup>11</sup> Thus the distinction between trust and reliance that figures so prominently in the philosophical literature is a distinction with application well beyond the interpersonal realm. Theories of trust that distinguish trust from reliance by giving an essential role to interpersonal attitudes like goodwill or resentment are being too restrictive.

<sup>12</sup> For applications of this methodology to promises with rather different results, see Scanlon (2003: 282–3) and Owens (2012: Chapter 6).

proposition asserted and it also explains why you can trust an assertion only by believing it, for you can learn the truth of what you are being told only by believing it.

In this paper I focus on promises.<sup>13</sup> Some writers have tried to explain the nature and value of a promise by treating a promise as ‘an invitation to trust’ the promisor. On this view, the normative significance of a promise depends on the normative significance of the attitude of trust that it invites.<sup>14</sup> Though a normal promise may well involve an invitation to trust, I doubt this observation explains the binding force of a promise.<sup>15</sup> To account for that we must move in the opposite direction: first identify the distinctive value of a promise and then deduce what trust in a promise must involve. We can agree that promises generally invite trust (and would be pointless unless they were sometimes trusted) without seeking to explain the point of a promise by reference to some promise-independent notion of trust, for it is impossible to understand what trust in a promise amounts to other than by reference to the distinctive value of a promise.

In the philosophical literature we find no consensus about what is involved in trusting either a promise or an assertion and I suspect these disagreements can be traced to differing views of the distinctive value these speech acts.<sup>16</sup> I shall rehearse several hypotheses about the distinctive value of promise but I’m not out to decide between them, only to use them to explain disagreements about trust in a promise. It may be that one hypothesis is correct and the others are wrong. It may also be that promises have more than one form of distinctive value and so there is more than one way of really trusting them. The point remains: our understanding of the nature and value of trust in a promise derives from our understanding of the distinctive value of a promise rather than that of trust.

## 2. Trust in a Promise: Social Coordination

A promise is made when the promisor communicates the intention of hereby undertaking an obligation to perform, a mouthful I’ll abbreviate by saying that promises are

<sup>13</sup> Owens (2012) offers a theory of promissory obligation that makes no mention of trust in a promise. In this paper I’m seeking to repair that omission whilst also giving the reason for it.

<sup>14</sup> For two accounts along these lines, see Friedrich and Southwood (2009: 282) and Pink (2009: 408). The explanatory burden here is thrown on the notion of trust and so one looks for an account of the nature of interpersonal trust that can reveal the basis of a specifically promissory obligation. Friedrich and Southwood (2009: 280) tell us that trust in a promise involves ‘a certain faith or optimism in the promisor’s character that the promisor will perform some action that is of importance to the promisee’. This formulation raises all of the questions to be addressed in what follows: Does ‘faith or optimism’ involve belief in or reliance on future performance? Can it be based solely on evidence of future performance? And what aspects of the promisor’s character and motivational psychology constitute suitable objects of trust?

<sup>15</sup> There are (binding) promises which do not invite trust in themselves but this involves adopting a perverse attitude to the value of the promise. See Owens (2012: 201–2).

<sup>16</sup> Were there a consensus both about the value of a promise and about the value of an assertion, this methodological point could instead be made by contrasting what is involved in trusting a promise with what is involved in trusting an assertion.

made by *declaration*. The promisor's declared intention to bind themselves is not enough to create the promissory bond, for the potential promisee must both understand what is happening and react to the promisor's offer—they must communicate the intention of hereby accepting the offer (Reid 2010: 336). In contexts where there is a default assumption that such offers will be accepted, silence is enough to effect this further declaration. In other contexts, the promisee may need to be more explicit. None of these claims about promise are uncontroversial, but for my purposes their truth is not crucial. I am using them to illustrate a certain methodological moral about the connection between the value of a promise and trust in a promise, a moral that does not depend on the particular model of promissory obligation we adopt.<sup>17</sup>

Turning now to *trust* in a promise, this also requires some form of uptake on the part of the promisee.<sup>18</sup> There are at least three possibilities:

- (1) *Trust is Acceptance*: You trust a promise just by accepting it (a speech act), thereby ensuring that it binds the promisor.
- (2) *Trust is Reliance*: You trust a promise by relying on the promise, by (at least) intending to behave as if it will be kept.
- (3) *Trust is Expectation*: You trust a promise where you believe that it will be kept.

These three forms of uptake, though distinct, often become entwined in practice but my present interest is in the differences between them.

My brother is in a generous mood and promises me an expensive birthday present. I wouldn't mind the present so I accept his offer whilst strongly suspecting that the present will fail to materialize. Here we have acceptance without either reliance on or expectation of performance. Acceptance can also be joined to the former without the latter. Though much of the time I rely on promises because I believe they will be kept, I sometimes have reason to rely on someone's promise even when I'm far from convinced that they will perform. Perhaps there is some reason to think they'll come through, enough to justify behaving as if they will (the stakes are low) but not enough to convince me that they will. Perhaps I have some reason to rely on their promise other than my conviction that they'll perform (e.g. as a display of confidence in them). Here we have reliance without expectation of performance.<sup>19</sup>

So what is involved in trusting a promise: must you believe that the promisor will perform, must you at least be prepared to rely on their promise or is it enough to simply accept their promise, at least where acceptance alone ensures that the promise binds?

<sup>17</sup> These assumptions about promissory obligation are in fact quite widely shared and are defended in Owens (2012). An expectations theorist like Scanlon would reject them but I believe the expectations theory could also be used to illustrate our methodological morals.

<sup>18</sup> Austin says that an illocutionary act like a promise secures uptake whenever it 'brings about the understanding of the meaning and force of the locution' (Austin 1975: 117). (1)–(3) all involve uptake in that minimal sense but all three go beyond it.

<sup>19</sup> Perhaps one can also expect performance without being willing to rely. When the stakes are high enough I don't feel I can expose myself to the risk of your not performing even though I do believe that you will perform (Jones 1996: 24).

This is the *modal* question, it being a question about the modality of trust.<sup>20</sup> Is trust a speech act or an attitude and, if the latter, is that attitude cognitive like belief or conative like intention? On my methodological hypothesis, we can address this issue (and others that will soon arise) only by asking ourselves what the distinctive value of a promise is.

In the eyes of many (Hume for example) the function of a promise is to facilitate social coordination by giving us all a way of getting other people to rely on us. I want help with my harvest and so I promise to help you with yours if you help me with mine (Hume 1978: 520–1). According to Hume's *social coordination hypothesis*, the key to realizing the coordinative value of a promise and thus to trust in a promise is reliance. Provided both parties rely on the compact between us, social coordination will be achieved regardless of whether each believes the other will perform. On this view mere acceptance need involve no trust in a promise since it need involve no tendency to act in reliance on the promise and so no tendency to coordinate my behaviour with the promisor's in a beneficial fashion. I may accept my brother's promise of a fine birthday present but I do not trust it. Thus the modal question is resolved: to trust a promise is to rely on it (i.e. to adopt a certain policy towards it) and I can rely on you to help me with my harvest without being convinced that you will. Here I trust your promise because I (intend to) behave in a way that realizes the distinctive coordinative value of the promise. Provided you adopt the same attitude to my promise, successful coordination will be achieved.<sup>21</sup>

Some writers do not agree that trust in a promise turns on reliance. They complain that simply to behave as if someone will perform without believing that they will is not really to trust their promise. On the other hand we do say 'You'll *have* to trust me, you have no alternative' or 'Being unsure whether you'd do it, I just had to trust you' and so forth. One might close this discussion down by multiplying senses of 'trust' or else conclude that there is no real disagreement here because a promise can be used either as an object of reliance or as a source of knowledge about what others are going to do and so trust in a promise can take more than one form. Each of these responses concedes that the dispute is verbal but, if the social coordination hypothesis is correct, we can resolve the matter in a theoretically motivated fashion: reliance is the attitude that realizes the distinctive value of a promise.

Let's now turn to a second question about the content of the attitude of trust. When trust in a promise is at issue, this becomes a question about what the motivational psychology of the trusted party must be like for the trust in them to be well placed;

<sup>20</sup> For opposing views on this question, see Holton (1994: 68–9), Hieronymi (2008: 216–19 and 227–31) and McMyler (2011: 131–41). On the whole these authors (and those cited in the notes that follow) do not focus on promises in particular. Rather they make general claims about the nature of trust, claims that I am applying to the case of trust in a promise.

<sup>21</sup> The possibility of relying on a promise without being convinced that it will be fulfilled ensures that we can *decide* to trust a promise; 'I wasn't sure you'd perform but I decided to trust you anyway' makes perfect sense where trust involves only reliance. By contrast it is doubtful whether we can decide to believe that the promisor will perform.

hence I dub this *the motivational question*. Promises are usually extracted where people feel they can't otherwise rely on the promisor to do the thing they are promising to do. If you know that the guy will mow your lawn anyway out of personal affection or because he loves mowing lawns, you probably won't seek a promise from him: promise is redundant, at least as a device of social coordination. Given that the promisor binds themselves by declaration, it is natural to assume that the function of a promise is to achieve social coordination in a special way, namely by imposing an obligation to perform on the promisor (Hume 1978: 518). If so, this fact should be reflected in our account of what it is to trust a promise, of how one must react to the promise in order to realize its distinctive value.

Here is one way of answering our motivational question: to trust a promise involves relying on the promise *because* one assumes that the promisor knows they are obliged to perform and is more likely to do so in the light of this fact. To trust a promise is not just to rely on the promise's being fulfilled for whatever reason; it is to rely on the promisor's conscientiousness in particular, making the assumption that conscientiousness is available at least as a back-up motive.<sup>22</sup> If one thinks the promisor will perform *only* from some unrelated motive like fear or favour, one is not trusting their promise (as opposed to trusting *that* they will do what they promised to do) because one is not reacting to the promise in a way that realizes its distinctive value, namely as an indicator of dutiful performance. And when the promisor breaches their promise they are not merely letting you down (like your car), they are *betraying* your trust by disregarding their obligations towards you.<sup>23</sup>

Not everyone agrees with Hume that the trusting promisee trusts the promisor to keep their promise out of fidelity to their promise. Some writers employ an obligation-independent notion of trust and argue that a promise involves an invitation to trust in *that* sense. On this conception of trust, neither giving one's own word nor accepting the word of another need involve thoughts of obligation. True, a promise usually binds the promisor to performance but on this view the obligation to keep a promise is a by-product of the relationship of trust:

The question when do gratuitous promises oblige comes down to this: under what conditions does inviting someone to trust you to do something oblige you to do it? And restating the

<sup>22</sup> Suppose the promisor regards their promise as a kind of vow—one they owe it to themselves to keep—rather than as a commitment, breach of which would wrong the promisee. Could the promisee trust this promise on the grounds that the promisor feels obliged to keep it?

<sup>23</sup> Pink observes that when someone breaches a promise, the promisee doesn't just resent being deprived of what they are owed, they also feel hurt at having been 'trifled with' (Pink 2009: 393 and 412). I resent it when, on a train, I leave some valuable exposed whilst using the restroom and it is gone when I return. Here I trusted my fellow passengers and my trust was *betrayed* but since none of them invited my trust, they did not *abuse* my trust and Pink's special sense of hurt is out of place (O'Neill 2012: 318–25). My trust would have been abused had one of them promised to take care of the valuable for me. A normal promise involves an invitation to trust and so breach of a promise normally involves an abuse of trust. This is so even when the promisee places no reliance on the promise provided acceptance alone can be a form of trust (see Section 3).

question in these terms really is explanatory. . . . For the idea of trustworthiness gives us independent purchase on when gratuitous promises give rise to obligations and when they do not. (Pink 2009: 411)

So what is this obligation-independent notion of trust?

It has been said that trusting someone to do something involves relying on them to be motivated to do the relevant thing at least in part because they know that you are relying on them to do it (Jones 1996: 8; Pettit 1995: 205–6). On this view an invitation to trust is an invitation to rely on me because I will be responsive to your reliance. Let's call this *interactive reliance*.<sup>24</sup> In the definition of interactive reliance, there is no mention of why the trustor supposes that their reliance on the trustee will move the trustee to perform. They might be assuming that the trustee will feel obliged to perform, but equally they might be assuming that the trustee feels a certain goodwill towards them or else wants to cultivate a reputation for reliability. Thus two people can engage in successful interactive reliance without either party imagining that obligation is involved.<sup>25</sup> Can we construct an obligation-independent notion of trust in a promise by understanding an invitation to trust as an invitation to interactive reliance?

Conceiving of trust in a promise as involving interactive reliance might seem a natural concomitant of the idea that a promise is a device of social coordination but I doubt an adherent of the social coordination hypothesis should endorse it. One can have successful interactive reliance (as I have defined it) in a situation where there is no promise at all and so talk of interactive reliance does not identify the special way in which a *promise* achieves social coordination. Suppose I ask you whether you are going to a certain conference with a view to deciding whether to go myself. You reply 'Well, as things now stand, I predict that I'll go but I can't promise you that I will go'. Perhaps you are more inclined to go precisely because we have had this conversation (you don't like to disappoint) and I may rely on that fact when I decide whether to go. Knowing this you *might* also feel obliged to warn me should you change your mind and so we may end up coordinating our plans in part by means of interactive reliance. But since I can't sensibly imagine that our conversation obliges you to actually go, I can't trust you to go out of fidelity to your word.<sup>26</sup> Here social coordination isn't achieved in the special way that a promise achieves social coordination.

I conclude that an obligation-independent conception of trust can't help to elucidate trust in a promise, but we might modify our definition of interactive reliance to include the idea that the trustee feels obliged to fulfil their promise only because the trustor is relying on them. Does trust in a promise involve interactive reliance so understood?

<sup>24</sup> This phrase comes from Pettit but he may mean something more by it than what is contained in my definition. Faulkner (2007: 881) argues that trust involves something like interactive reliance, though he adds that resentment would be apt should the trust be misplaced.

<sup>25</sup> This may be what is going on in Pink's doctor example (Pink 2009: 394–5).

<sup>26</sup> 'A purpose is no contract, even when it is declared to the person for whose benefit it is intended. I may say to a man that I intend to do such a thing for your benefit, but I come under no engagement' (Reid 2010: 336).

I'm unsure but I suspect not. Clearly you can rely on someone to do something out of fidelity to their promise without imagining that your reliance will move them to keep the promise: perhaps they don't know that you are relying on them and don't need to know in order to feel bound to keep the promise, or perhaps they suspect you don't trust them and are seeking to demonstrate their conscientiousness. In any case, the question for us is whether such reliance amounts to trusting the promise.<sup>27</sup> Since the social coordination hypothesis does not obviously settle the matter, the issue may be verbal. We'll revisit this point in the next section but for now I'll proceed on the assumption that you can trust someone's promise without assuming that they will be moved by your trust.

We can sum up our first conception of trust in a promise, based on Hume's social coordination hypothesis, as follows:

- (A) You trust someone to keep their promise iff you rely on them to perform because they know they are obliged to perform.

Note that, on this conception, there is no concern with how the promisee comes to rely on the promisor's conscientiousness. For all (A) says the promisee might credit the promisor's assurances only because they have conducted a careful analysis of the promisor's past behaviour and, on that basis, have concluded that the promisor is conscientious. This observation raises a third and final issue about trust in a promise that I've dubbed *the question of justification*.

This question concerns how the promisee's reliance on the promisor's conscientiousness must be justified for that reliance to count as trust in their promise. Those who think that trust involves belief in performance must insist that reasons for trust include whatever evidential (or other) reasons are required to justify the belief that the promisor will perform. Having allowed that a promise can serve its characteristic purpose without convincing the promisee that the promisor will perform, we can afford to be more liberal on the point. An advocate of the social coordination hypothesis should permit my trust in someone's promise to be based on anything that makes reliance reasonable; that includes both evidence of the promisor's conscientiousness and also the non-evidential value of reliance. Thus I might rely on my brother's conscientiousness, even though I do not myself regard him as conscientious simply because my mother has asked me to. On the social coordination hypothesis I am here trusting my brother even though I do not see him as trustworthy because I am adopting an attitude towards him which will realize the distinctive value of his promise should he turn out to be conscientious.

<sup>27</sup> O'Neill (2012: 318–25) might describe this as a case in which the promisor is trusted not to betray trust rather than not to abuse it. Nickel (2012: 307–9) makes the point that you trust someone in relying on their conscientiousness even if you don't suppose that they will come through *because of your trust*, but Nickel appears to deny that this can happen in the case of a promissory obligation. I agree that a promise generally invites trust but the trust it invites may just involve reliance on the conscientiousness of the promisor.

There are now two answers to the question of justification on the table but we are not yet done. For some writers, to insist on having evidence sufficient to justify the belief that John will perform would be to manifest a *lack* of trust in John: trust actually *excludes* such confidence. On this third view, a paradigm instance of trust is where a prospective reformer is asked to employ a man convicted of theft in their shop (Holton 1994: 63).<sup>28</sup> Here the reformer is in no position to know whether the convict will steal but might well decide to trust him nevertheless. As Faulkner puts it:

if the reformer is willing to depend *only* given this belief and good reason... then she does not really trust. Too thorough an assessment of the risk is inimical to trust. (Faulkner 2007: 897)<sup>29</sup>

Again there is a temptation to respond to such examples by multiplying senses or else by selectively highlighting those bits of ordinary talk about trust that appear to favour one usage of 'trust' over another. Again the issue acquires theoretical significance only once we link it to developed views about the characteristic value of a specific object of trust like a promise. An advocate of the social coordination hypothesis will reject Faulkner's observation but, as we'll see in the next section, a different account of the value of a promise delivers a rather different answer to some and perhaps to all of our three questions.

### 3. Trust in a Promise: Cultivating Relationships

On the social coordination hypothesis, the value both of a promise and of trust in a promise are fundamentally instrumental; they render the future behaviour of others predictable and such predictability yields many further benefits. Alternatively it has been proposed that trust may be valued for its own sake because trusting relationships are to be valued for their own sake and not merely for the further benefits they provide. This opens up the possibility that the distinctive value of a promise is something to do with the non-instrumental value of these relationships.

Let's leave trust in a promise behind for a moment and consider the following example:

Suppose we are rock climbing together. I have a choice between taking your hand, or taking the rope. I might think each equally reliable; but I can have a reason for taking your hand that I do not have for taking the rope. In taking your hand, I trust you; in so doing our relationship moves a little further forward. This can itself be something I value. We need not imagine that you would be hurt if I chose the rope over your hand; you might be perfectly understanding of the needs of the neophyte climber. But our relationship would not progress. (Holton 1994: 69)

Here Holton takes the hand of his companion because he values having a climbing buddy and not just a source of physical support. Were the latter all that mattered to

<sup>28</sup> In other cases (close relatives for example) one may have conclusive evidence of reliability but one's trust in them cannot, on this third view, depend upon it.

<sup>29</sup> For an opposing view, see Pettit (1995: 207–8).

him, the rope might be a better buddy (and one that demands no return). Holton clearly expects that taking the hand rather than the rope will help to bring that valuable relationship into existence.

We might use the notion of interactive reliance to understand what is going on here. Perhaps what Holton values for its own sake is someone's holding him in part because Holton is relying on them to hold him. That does sound like a human connection that one might value for its own sake and one engendered by reliance. Furthermore this sort of reliance often creates obligations. By accepting his companion's hand, Holton ensures that his prospective climbing buddy is obliged to hold him and would be wronging him by letting him go (even if this would not lead to a serious fall). But, it might be thought, such obligations are a by-product of the real source of the relationship's value, namely interactive reliance.

At the outset, I characterized trust as something that would realize the distinctive value of the object of trust. Suppose that interactive reliance on another person is valuable for its own sake and constitutes the distinctive value of climbing buddy relationships. Given that trust realizes this value, we can answer the questions of justification and motivation as applied to Holton's example. The latter question concerned the trustee's motivation and in particular whether they should be moved to perform by the fact that the trustor is relying on them to perform and the answer is now a definite 'yes'. As to the justification for such trust, on the present view to be motivated to trust someone's offer of a hand simply by evidence of their reliability is not to be responsive to the distinctive value of what is on offer; rather it is to be responsive only to its value as an indicator of future performance and not to the non-instrumental value of interactive reliance. It is to treat the climbing buddy relationship as having purely instrumental value. That is why one might (like Faulkner) regard such caution as untrusting. Once more, we see how differing views about the characteristic value (instrumental or non-instrumental) of relationships like being someone's climbing buddy generate different views about what constitutes a trusting response to your companion's offer of a hand.

Holton's example need involve no promise. The parties may realize the normative significance of the offer and acceptance of a hand (i.e. the obligations it will entail) but they need not intend to change the normative situation, for they might have no interest in who is obliged to do what. How can we generalize from Holton's example to the case of promise where the focus is on the creation of an obligation? In the last section, I argued that interactive reliance is insufficient for trust in a promise: the trusting promisee must also assume that the motive of duty is in play. Given this, we could adopt something like the following conception of trust in a promise:

- (B) You trust someone to keep their promise iff (i) you rely on them to perform because they know they are obliged to perform and (ii) you so rely on them in part because of the non-instrumental value of the obligation to perform.

On this conception, part of the value of a trusting relationship lies in the very obligations that constitute such a relationship. Call this the *normative interest hypothesis*. (B) uses the notion of reliance to formulate that hypothesis.

Trusting relationships are human connections with a certain normative character, namely connections that involve a network of reciprocal rights and obligations. Friends, neighbours, family members, and climbing buddies are embroiled in such relationships. I suggest that the rights and obligations constitutive of these relationships are not mere by-products of the features that give those relationships their value. Rather, they contribute a deontic element to that value. Furthermore, this deontic value is at least part of the distinctive value of such a relationship, of the value that explains the binding force of its constituent obligations. Friends rightly value the bonds of loyalty for their own sake and they are bound to one another in part because being so bound makes their lives go better (Owens 2012: Chapter 4). To generalize, the obligations that constitute relationships of trust are (a) valuable for their own sake at least in the context of the relevant relationship and (b) bind the parties in part *because* they are valuable for their own sake.

The offer and acceptance of a hand in Holton's example probably aren't declarations that constitute the making and acceptance of a promise but, on our new conception of the value of a trusting relationship, they may have much the same rationale and effect as clear cases of promise. Here is such a case. Someone has moved in next door and I wish to cultivate a neighbourly relationship with them, so I accept their offer to water my plants while I am on holiday, knowing that I must make a similar offer when they go away. I do this even though I could easily move the plants to my son's house for the duration because I wish to build a network of rights and obligations constitutive of a relationship with a certain sort of non-instrumental value. Where the conventions of neighbourly behaviour don't already tell us how to behave, the way to do this is via explicit agreements (Raz 1986: 173–6). Note that if our relationship develops as envisaged, these neighbourly obligations will soon cease to be obligations based on a dateable promise and will become part of how we expect our neighbour to behave towards us in virtue of the nature of our relationship (Owens 2012: 106–7), but in the meantime we use our power of promise to create obligations that we value (or envisage coming to value) for their own sake.<sup>30</sup>

(B) tells us that trust in a promise is reliance motivated at least in part by the non-instrumental value of the promissory obligation. This answers the question of justification. Someone who relies on the promise simply because there is strong evidence that it will be kept is treating the promise as having purely instrumental (i.e. evidential) value and so is not trusting it.<sup>31</sup> What of the modal question? Here we encounter a

<sup>30</sup> Here I'm focused on promises made in the context of a developing relationship whose value depends on that context but I don't mean to imply that obligation can't matter to people (and for its own sake outside that context. We can both make binding promises and trust them whenever it makes sense to value an obligation (and the power to impose it). See Owens (2012: Chapter 6).

<sup>31</sup> There are other ways of treating the promissory obligation as having purely instrumental value. I might rely on your promise out of politeness or because my mother told me to and without considering

potential problem. If valuing the obligation is the essence of trusting a promise (rather than valuing the promised act) it looks as if the basic form of trust should be acceptance of the promise rather than reliance on the promisor's behaving as promised, for the promise binds provided it is accepted and regardless of whether the promisee relies on it. If so, the normative interest hypothesis should be formulated using the notion of acceptance rather than reliance.

But can simply accepting a promise really be a way of trusting the promise? There are cases and cases. Obligations and the promises that create them are frequently valued in the context of a developing relationship and where this is so I can often realize the value in question only by relying on the relevant promises. I value my obligation to water my neighbour's plants as one aspect of an ongoing relationship with my neighbour, as part of a network of rights and obligations (and of the habits and customs which embody them). That network will not come into existence unless the promises we make to one another are not merely accepted but also relied upon, for our relationship will not develop in the envisaged way if, having accepted the neighbour's promise, I panic and transfer the plants to my son's house. Here my neighbour is no longer obliged to water my plants.<sup>32</sup> In failing to rely, I decline to make myself vulnerable to being wronged by my neighbour (in respect of non-performance) and so I fail to create the debt that will be the occasion of my later promising to water their plants whilst they are away. The desired connection between us never materializes and the promissory obligation loses its rationale.

But in other cases (e.g. my brother's promise of a present) there may be nothing I can do or fail to do that would count as reliance on the promise beyond simply accepting it. Here acceptance makes me vulnerable to being wronged and such acceptance all by itself constitutes trust in a promise (provided that acceptance is motivated by my valuing such vulnerability for its own sake). If I accept your promise just to please my mother and without any real interest in whether you are bound to me or not, I am still wronged by your breach but it would be odd to describe you as betraying my trust or even as letting me down. On the other hand, if I accept your promise because I want to create a relationship of mutual vulnerability between us, because I want to make room for the possibility of fidelity and betrayal, then acceptance even in the absence of reliance on or expectation of performance can indeed amount to trusting your promise.

I've rather fallen out with a friend who has stood me up once too often. They solemnly promise to help me move house and I accept their promise as a way of giving them the opportunity to restore relations between us, even though I have an open mind about whether they'll show up, and things are arranged so that all will go smoothly whether or not they appear. On the present view of promise, I do trust them

whether your promise is a good predictor of performance. In these cases also I am relying on the promise without trusting it (on the present view) because my reliance is not based on my valuing the promissory obligation for its own sake.

<sup>32</sup> This is so in virtue of the impossibility of performance and regardless of whether I have explicitly released the neighbour from their promise.

and they can betray my trust even though I place little credence in or reliance upon what they say. After all, I've opened myself up to being wronged. Here the fact that they would wrong me by not showing up has a significance of its own *and* I accept their promise because I wish to give their behaviour that significance, to make myself vulnerable to them in that way.<sup>33</sup>

Trust in a promise is an attitude that realizes the distinctive value of the promise, a value that, on the normative interest hypothesis, is the (non-instrumental) value of the promissory obligation. Acceptance and reliance can each play this role on different occasions and so both can constitute trust in a promise. Thus the normative interest hypothesis and the social coordination hypothesis differ in their answers to both the modal question and the justificational question. Our conception of trust in a promise once more depends on our conception of the value of trust's object, namely the promise.

#### 4. Conclusion

The recent debate about the nature of trust is intricate and inconclusive because its participants see themselves as analysing a distinctive psychological attitude, an attitude about whose nature and value they are all disagreeing. I suggest that the tangle makes more sense if we regard it from the other end, starting with the objects of trust. With so many different objects of trust in play, and so much disagreement about their nature and value, an inconclusive debate is only to be expected. We made this point by focusing on trust in a promise and I argued that one's view of that should depend on one's conception of the distinctive value of a promise.

I did not seek to establish the correctness of any particular conception of the value of a promise. Indeed we didn't even consider one of the most influential accounts of this matter, namely the expectations theory (Scanlon 2003). According to that theory, the point of a promise is to provide the promisee with knowledge of what the promisor is going to do. For Scanlon, trust in a promise involves the belief that the promisor will perform, a belief for which the promisee must have sufficient evidence. On this conception of the value of a promise our answer to the modal and justificational questions for promise will likely be similar to those applicable to assertion, whilst on both the social coordination and the normative interest hypothesis these answers are going to be very different.

Elsewhere I argue that neither the social coordination hypothesis nor the expectational account provides an adequate conception of the distinctive value of a promise and I endorse the rival normative interest hypothesis (Owens 2012). Nothing I say here

<sup>33</sup> Here the obligation has value only in the context of the restored friendship and the friendship will be restored only if the former friend fulfills their promise. So they are worthy of my trust only if they turn out to be reliable but trusting them involves no more than accepting their promise. Thanks to Collin O'Neil for forcing me to clarify this.

rests on that claim. What this paper seeks to establish is that none of our three questions about trust can be settled by applying some general theory of trust to the case of promise; rather they must be approached from the opposite direction by first formulating a view of the value of a promise. Trust in someone's assertions or trust in their courage should be tackled in the same fashion. There is no general attitude of trust fitted to play a foundational role in our social theory. Trust is as various as the objects of trust.<sup>34</sup>

## References

- Austin, J. (1975) *How To Do Things With Words*, 2nd edn. (Cambridge, MA: Harvard University Press).
- Baier, A. (1986) Trust and Anti-Trust, *Ethics* 96: 231–60.
- Faulkner, P. (2007) On Telling and Trusting, *Mind* 116: 875–902.
- Friedrich D. and N. Southwood (2009) 'Promises and Trust', in *Promises and Agreements*, ed. H. Sheinman (Oxford: Oxford University Press): 277–94.
- Hieronymi, P. (2008) The Reasons of Trust, *Australasian Journal of Philosophy* 86: 213–36.
- Holton, R. (1994) Deciding to Trust, Coming to Believe, *Australasian Journal of Philosophy* 72: 63–76.
- Hume, D. (1978) *Treatise on Human Nature* (Oxford: Oxford University Press).
- Jones, K. (1996) Trust as an Affective Attitude, *Ethics* 107: 4–25.
- McMyler, B. (2011) *Testimony, Trust, and Authority* (Oxford: Oxford University Press).
- Nickel, P. (2012) Trust and Testimony, *Pacific Philosophical Quarterly* 93: 301–16.
- O'Neil, C. (2012) Lying, Trust and Gratitude, *Philosophy and Public Affairs* 40: 301–33.
- Owens, D. (2012) *Shaping the Normative Landscape* (Oxford: Oxford University Press).
- Pettit, P. (1995) The Cunning of Trust. *Philosophy and Public Affairs* 24: 202–25.
- Pink, T. (2009) Promising and Obligation, *Philosophical Perspectives* 23: 389–420.
- Raz, J. (1986) *The Morality of Freedom* (Oxford: Oxford University Press).
- Raz, J. (2001) *Value, Respect, and Attachment* (Cambridge: Cambridge University Press).
- Reid, T. (2010) *Essays on the Active Powers of Man*, ed. K. Haakonssen and J. Harris (Edinburgh: Edinburgh University Press).
- Scanlon, T. (2003) Thickness and Theory, *Journal of Philosophy* 100: 275–87.

<sup>34</sup> Many thanks to Robert Stern, Alison Hills, Paul Faulkner, Tom Simpson, Jessica Moss, Sharon Street, David Velleman, Matty Silverstein, Jorah Dannenberg, Nic Bommarito, and especially to Collin O'Neil for comments.