

Princeton University Press

The Cunning of Trust

Author(s): Philip Pettit

Source: *Philosophy and Public Affairs*, Vol. 24, No. 3 (Summer, 1995), pp. 202-225

Published by: Formerly published by Princeton University Press

Stable URL: <http://www.jstor.org/stable/2961900>

Accessed: 11/11/2008 09:33

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=pup>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



Princeton University Press is collaborating with JSTOR to digitize, preserve and extend access to *Philosophy and Public Affairs*.

<http://www.jstor.org>

Trust materializes reliably among people to the extent that they have beliefs about one another that make trust a sensible attitude to adopt. And trust reliably survives among people to the extent that those beliefs prove to be correct. Trustors identify reasons to trust others and trustees show that those reasons are good reasons: the trust which they support is generally not disappointed.

It is important to be clear about the reasons, in particular the good reasons, why people might invest trust in one another. For a society where people are disposed to be trusting, and where their trust is generally well placed,¹ is almost certain to work more harmoniously and fruitfully than a society where trust fails to appear or spread.² If we are not clear about the good reasons why people might trust one another, we are in danger of designing institutions that will reduce trust or even drive it out.

This article is a contribution to the project of understanding the reasons—and, potentially, the good reasons—why people might trust one another. I discuss some more or less standard reasons why people trust

I thank John and Val Braithwaite for getting me interested in trust and for giving me some insight into it. I am grateful to a large number of people for the incisive comments that I received when an earlier version of the paper was presented at a workshop on trust in the Australian National University, and at seminars at the University of Colorado, Boulder, and Stanford University. I cannot hope to name them all. I was led to make a number of significant changes by the comments of the editors of *Philosophy & Public Affairs*.

1. Russell Hardin, "The Street-Level Epistemology of Trust," *Politics and Society* 21 (1993): 505–29.

2. Diego Gambetta, ed., *Trust: Making and Breaking Cooperative Relations* (Oxford: Blackwell, 1988); Robert D. Putnam, *Making Democracy Work: Civic Traditions in Modern Italy* (Princeton: Princeton University Press, 1993); Francis Fukuyama, *Trust* (New York: Basic Books, 1995).

one another, describing these as reasons of trustworthiness, and then I argue that so far as reasons of trustworthiness are recognized in a culture, there is also a further sort of reason available to support trust. This is a consideration to the effect that even where others are not independently known to be trustworthy in the standard way—even where they are not independently known to have the desirable traits associated with trustworthiness—they can be presumed to be responsive to acts of trust.

The trust-responsiveness that I identify constitutes a disposition to prove reliable under the trust of others, and in this respect it is similar to trustworthiness; both are forms of trust-reliability. But whereas the different forms of trustworthiness represent traits that all see as desirable, at least in certain respects—hence the worthiness—the trust-responsiveness that I have in mind is not a trait that many will be proud to acknowledge in themselves. It is the desire for the good opinion of others and it counts by most peoples' lights, not as a desirable feature for which they need to strive, but rather as a disposition—a neutral or even shameful disposition—that it is hard to shed. The fact that it can be supported by such a disposition shows a certain cunning on the part of trust. Trustors do not have to depend on the more or less admirable trustworthiness of others; they can also hope to exploit the relatively base desire to be well considered.

The article is in five sections. First, I offer a characterization of the sort of trust with which I am concerned. Next I look at some trustworthiness reasons that can serve to sustain such trust; these are reasons associated with the ascription of traits like loyalty and virtue and prudence. In the third section I make the case for a further, trust-responsiveness reason to trust: the reason associated with people's love of regard or standing in the eyes of others. In the fourth section I emphasize the importance of this reason by drawing attention to some aspects of trust that it helps to explain. And then in the fifth section I show that it is important to understand this reason to trust if we are not to make mistakes in institutional design: The desire for the good opinion of others will facilitate trust only under certain institutional conditions.

I. THE CHARACTERIZATION OF TRUST

The word 'trust' is used in relation to a great number of things. The word may be used in connection with relying on natural phenomena as well

as in connection with relying on people. When it is used of relying on people, it may apply to relying on them to have certain skills or capacities as well as to relying on them to act in certain ways. And when it is used of relying on people to act in certain ways, as it is used in most discussions of trust, and as I shall use it here, it may apply to any of three distinct phenomena.

The most general usage of the word in this connection would equate trust with confidence that other people will treat you reasonably well: confidence that they will not waylay or cheat you, for example. We speak in this sense of trusting our fellow citizens or trusting the institutions under which we live. A somewhat less general usage would link it with confidence that other people are reliable under certain tests: they will treat you well, in the event of your placing yourself in their hands. We speak in this sense of trusting the police or trusting the courts. A third usage, more specific still, would associate trust, not with a detached confidence that people are reliable under such tests, but with putting that confidence to the test: with actually placing yourself in the hands of others. We speak in this sense of trusting the police or the courts when we actually call on the police for help or take a complaint to the courts.

These three phenomena—these three forms of reliance on people's behavioral dispositions—are all of importance in social life, and all of them attract and deserve the name of 'trust.' But my concern in this article is not with such phenomena generally, only with the sort of case where you place yourself in the hands of another: only with the case of active reliance, as we might call it. In this case you rely on others to the extent of making yourself vulnerable to them, voluntarily or under the force of circumstance. The most salient example is where you rely in your own individual right on another individual person. But in other cases you may rely on a certain agent in tandem with other individuals, and you may rely on a corporate or collective agent that itself involves a number of people.

Active reliance is still too broad a category, however, to capture the object of my concern in this article. The reliance with which I shall be concerned is not just active, but interactive. It is a form of reliance that is distinctively trusting, in a perfectly ordinary sense of that term.

Suppose I am driving into a city that I do not know and I wish to get to the town center. I see a bus and, knowing the pattern on which bus

routes are generally organized, I decide to rely on the bus driver to lead me to the center. This is a straightforward case of active reliance. I rely on the driver to behave in a certain way in the sense that I build my own plans around the assumption that the driver will behave in that way. I assume the driver is so motivated and so informed that he will behave appropriately; or I assume that that is a good bet, or as good a bet as any other available to me.³ Assuming this, I give over control of certain of my fortunes—or of the fortunes of those with whom I identify—to the driver; I bind the welfare of me or mine to his performance.

But the reliance that this example illustrates assumes a more specific and interesting form if it becomes interactive as well as active. Suppose that I worry about what the bus driver will think about a car that stops every time the bus stops and that follows the bus faithfully on its route. This may lead me to get out at a bus stop and let the driver know that I am relying on him to lead me to the center and that that is why I am staying behind the bus. If I do that, then my reliance becomes manifest: The driver knows that I am relying on him and knows that I am aware that he knows that. Perhaps the reliance even becomes a matter of common knowledge, with each of us being aware of the reliance, each being aware of this awareness, each being aware of that higher-order awareness, and so on.⁴

The object of my concern here is interactive reliance of this kind, not just active reliance. But the object of my concern is not interactive reliance in general, only a sort that can be characterized as distinctively trusting. What I shall be referring to in speaking of trust is this trusting, interactive reliance, not just reliance of any old kind.

The distinction between the two sorts of interactive reliance can be brought out, once again, by reference to the bus-driver example. When I let the driver know that I am relying on him to get me to the city center, I may do this in either of two minds. I may not expect that the driver cares in any way for my welfare; I may even think that the driver is

3. Richard Holton, "Deciding to Trust, Coming to Believe," *Australasian Journal of Philosophy* 72 (1994): 65–66.

4. We might be content with a weaker account of common belief: say, one which requires that each believes *p*, that each believes that each believes it, that no one disbelieves that each believes this, that no one disbelieves that such disbelief is lacking, and so on (David Lewis, *Convention* [Cambridge, Mass: Harvard University Press, 1969], pp. 165–66). A hierarchy of such disbelief is easy to live with, because it requires only the absence of certain higher-order, complex beliefs, not their presence and proliferation

malevolent, on the ground that bus drivers generally take pleasure in frustrating members of the public. Or I may expect that the driver will be positively moved by seeing that I have made myself vulnerable and will be motivated all the more strongly to do that which I am relying on them to do: will be motivated all the more strongly to prove reliable.

In the first case, the interactive reliance that I display is not particularly trusting. I rely on the bus driver, because I know how bus routes are laid out in cities like *this*, or perhaps because I see 'City Center' displayed on the bus. I rely on the bus driver, despite my thinking that he is indifferently or even malevolently disposed toward me. I rely on him, solely because I reckon that he is constrained to behave in the required fashion.

In the second case, which is the one that interests us here, the interactive reliance that I display is distinctively trusting. I see the driver as someone who is well-disposed toward me, whether in my individual right or as a member of the public, and I believe that my manifesting reliance will strengthen or reinforce his existing reasons to do that which I rely on him to do.⁵ For whatever reason, I assume the attitude of a trusting individual.

What can it mean to believe that the bus driver's reasons for acting in the required way are strengthened or reinforced, if I already believe, as well I may, that there is little or no possibility of his letting me down: If I think that he is bound to go to the city-center destination that is advertised on the bus? I already believe in such a case that the driver's utility for getting to the city center is higher than the utility he attaches to going anywhere else. But I will be trusting in my attitude toward the driver if I also believe, on the grounds of his being well disposed, that the utility he attaches to getting to the city center increases with the recognition that getting there will serve my purposes.

Interactive, trusting reliance can be characterized, then, by three clauses. One person relies in this way on another to the extent that:

1. He or she relies on another to do something, A;
2. this reliance is manifest to the other; and
3. the first person expects the second to be well disposed and to attach a greater utility to doing A for the fact that it represents a way of proving reliable.

5. Annette Baier, "Trust and Antitrust," *Ethics* 96 (1986): 231–60.

This account needs some qualifications to cover cases involving corporate or collective agents. For example, if the trustor is a collectivity of some kind, then it may only be the reliance of the group, not of any particular individual, that has to be manifest to the trustee and that has to motivate the trustee. But we need not concern ourselves with such details here. The conditions serve pretty well to identify the object of my concern in this article.

Interactive, trusting reliance, as I have stressed, is not the only thing that we use the word 'trust' for. When I focus on such reliance, I do not mean to suggest that it has any monopoly claims on the name of 'trust.' And when I identify the conditions under which such reliance is present, I do not mean to present them as conditions in the analysis of the concept of trust. Henceforth I shall use the word 'trust' only for interactive, trusting reliance, but I follow this practice just for reasons of convenience; I do not renege on any of the points emphasized here.

Why focus on this sort of trust, rather than on reliance or active reliance or even interactive reliance more generally? For one thing, it is always good practice to sharpen the object of concern in an exploration of this kind and only to look later for possibilities of generalization. But a second reason is that interactive, trusting reliance has certain normatively attractive features that make it particularly worth investigating. Where trust of this kind materializes and survives, people will take that as a token or proof of their being well disposed toward one another, so that the success of the trust should prove to be fruitful in other regards. Whatever the evaluative stance from which trust is viewed, that result is bound to present itself as, in general, a good thing.

In conclusion, a query. I have assumed that interactive, trusting reliance deserves the name of 'trust,' even while admitting that it is not the only deserver of that name. But someone may say that trusting always means taking a risk and that the account allows that I may trust someone to do something even when I have independent reasons to be sure that they will do it: this, as in the bus-driver case. Thus they may claim that the phenomenon I target does not strictly deserve to be called trust. "As virtually all writers on the subject agree, trust involves giving discretion to another to affect one's interests. This move is inherently subject to the risk that the other will abuse the power of discretion."⁶

6. Russell Hardin, "The Street-Level Epistemology of Trust," *Politics and Society* 21 (1993): 505–29, esp. 507.

The objection misfires, because it is surely plausible that I may trust someone of whose behavior I am independently assured. I may trust a friend to do A though, for any of a variety of reasons, I cannot imagine her doing anything other than A: the reason may be that the law requires that she do A, that doing A is a matter of virtue or honor, that she is indeed a very good friend, or whatever. My reliance on her will not lower the utility she attaches to doing A, as it would if she were ill disposed. Nor will it fail to raise that utility, as it would if she were indifferent. I trust her to the extent that I expect my reliance to strike a responsive chord—she is well disposed—and to raise the utility that she attaches to doing A.

The objection is probably inspired by an ambiguity in the notion of risk-taking. To trust someone in our sense may not always be to take a risk, in the sense of relying on that person to do something which you are not assured he will do. But it will always be to take a risk in another sense: It will always be to make yourself vulnerable to the other person in some measure, to put yourself in a position where it is possible for the other person, so far as that person is a free agent, to harm you or yours.⁷ I may run no probabilistic risk, as I see things, in relying on you to do A. But I must still recognize that you are a free agent and that my welfare is in your free hands.

II. MECHANISMS OF TRUSTWORTHINESS: LOYALTY, VIRTUE, AND PRUDENCE

There is no problem about why people should rely from time to time on others; this may be required for attaining their ends. And equally there is no problem about why they should make it manifest, if it is not manifest already, that they are relying on others in this way; they may often have no option but to make it manifest. But why should people believe that others are well disposed and that manifesting their reliance to another is likely to raise the utility that the other attaches to performing in the manner required?

To the extent that we find reasons why people actually hold a belief in what we may call the motivating efficacy of manifest reliance—in the efficacy, now in this situation, now in that—we will have revealed mech-

7. Richard Holton drew this point to my attention. I am particularly indebted to Geoff Brennan for discussion of the point.

anisms whereby trust is aroused amongst them. And to the extent that we find good reasons why they should hold such a belief we will have revealed mechanisms whereby trust is sustained and spread.

Three sorts of reasons are regularly associated with trust, and these do give us considerations that may be expected to generate trust—in particular, to generate a belief in the motivating efficacy of manifesting reliance—and to generate it fairly reliably. They are, respectively, reasons of loyalty, reasons of virtue, and reasons of prudence.

Suppose I believe that someone is a loving family member, a loyal friend, a devoted colleague, or whatever. This belief offers one ground on which I may expect that if I manifest the fact that I am relying on that person to do something then that person will be led to attach a greater utility to doing it. Loyalty—specifically, the trustor's belief in the loyalty of the trustee—offers a first mechanism whereby trust may be aroused and sustained.

Or suppose I believe that someone is virtuous: say, a god-fearing sort who can be relied upon to follow certain religious norms. This belief will offer a different ground for thinking that if I manifest the fact that I am relying on the person to do something then that will help motivate him to do what I require. The sort of virtue envisaged would make it difficult for the person to let down someone who depends on him in the manner associated with the act of reliance; it would represent the act of proving reliable as an act that virtue requires of him and would raise the utility that he attaches to it. And so I may expect that just by manifesting my reliance I can tap into that virtue and help to secure the sort of performance I want.

Or suppose I believe that someone is a prudent sort who will see the potential long-term rewards of maintaining a certain relationship: in particular, a relationship that requires her to prove responsive to certain acts of reliance on my part. This belief will offer a third ground for thinking that by manifesting the fact of relying on her to do something appropriate, I can actually motivate her to perform accordingly: I can alert her to the potential rewards of proving reliable and thereby maintaining the relationship—perhaps just a trading relationship—with me.

Not only can the mechanisms of loyalty, virtue, and prudence make it sensible for me to believe in the motivating efficacy of manifesting reliance, and make it sensible for me to trust the person in question in

a relevant domain. The mechanisms can also explain why trust builds on trust: why trust tends to grow with use, not diminish.⁸ For it should be clear that as I test and prove someone suitably loyal, suitably virtuous, or suitably prudent, I have reason to be reinforced in my disposition to put those mechanisms to the test in future acts of trust. Moreover, as I display a belief in the efficacy of loyalty or virtue or prudence, this should give the other person reason to assign to me a corresponding disposition not to let him down when he manifests similar acts of reliance; and so it should give him reason to invest more and more trust in me, as I invest more and more trust in him. Or at least it should do this so far as the relationship between us is saliently symmetrical: it is such that I cannot reasonably form certain expectations about the other's treatment of me without expecting him to form corresponding expectations about my treatment of him; and this, as a salient matter that each of us should come to recognize in common.

The mechanisms of loyalty, virtue, and prudence are not exclusive of one another. Indeed it should be clear that they are capable of reinforcing each other in supporting acts and relationships of trust. I have all the more reason to expect people to be motivated suitably by my manifesting a certain reliance, if I see them as susceptible, not just to loyalty, or virtue, or prudence, but to two or more of these traits at once. I may recognize that while loyalty or virtue will probably be the motor that leads them to perform as I rely on them to perform, for example, still I need not be worried about the possibility of that motor failing; even if it fails, the engine of prudence is there to take over the work of ensuring performance.⁹ Or I may see them as being motivated in a mixed fashion, with loyalty and virtue and prudence each having a shoulder at the wheel that controls their behavior.

I mention the possibility of these three mechanisms supporting one another, because many real-life examples of any one mechanism are likely to represent potential examples of the others too. Suppose a trustor expects ties of family loyalty to motivate a response to some act of reliance. That expectation may well be strengthened by the belief that

8. A. O. Hirschmann, "Against Parsimony: Three Ways of Complicating Some Categories of Economic Discourse," *American Economic Review Proceedings* 74 (1984): 88–96.

9. Philip Pettit, *The Common Mind: An Essay on Psychology, Society and Politics* (New York: Oxford University Press, 1993; reissue 1996), Chap. 5; Pettit, "The Virtual Reality of Homo Economicus," *Monist* 78 (1995): 308–29.

the trustee is virtuous and, failing the spontaneous firing of family affection, will be motivated by more austere, moralistic considerations not to let the trustor down. And equally the expectation may be strengthened by the belief that the trustee is prudent and, failing spontaneous loyalty or virtue, will recognize that letting the trustor down would cause them a net, long-term loss.

The mechanisms of loyalty, virtue, and prudence can be seen in combined operation in a nice example of trust relationships that has been analyzed by Avner Greif.¹⁰ This is the case of a widely dispersed network of medieval traders, all Jewish and all associated with the Maghrib: that is, with the western end of the Muslim world. The traders were successful, so the record goes, to the extent that they were able to maintain relationships of trust with one another and surmount contemporary difficulties associated with lack of mutual scrutiny and control, and an inability to enforce contracts legally. How was their trust sustained? Partly, by the mutual loyalty of a group identified as “our people, the Maghribis, the travelers.”¹¹ Partly, by the ability of the Maghribis each to advertise themselves, against the background of a common religion, as god-fearing and virtuous.¹² And partly—mainly, under Greif’s account—by the widespread perception among the group that their long-term, prudent interests were best served by a reputation for being trustworthy: “The agent cannot increase his lifetime utility by cheating.”¹³

The mechanisms that we have surveyed in this section are all fairly straightforward and salient. To be loyal or virtuous or even prudent is, in an obvious sense of the term, to be trustworthy. It is to be reliable under trust and to be reliable, in particular, because of possessing a desirable trait. The trait involved in each case can give a potential trustor reason to think that manifesting a certain reliance will prove motivating in the appropriate way: it will find a responsive chord in the trustee and it will raise the trustee’s utility for behaving in the fashion required. Believing that someone is loyal or virtuous or prudent in the appropriate way is just believing that they are trustworthy. And there is no mystery about how a belief in the trustworthiness of a trustee can serve as a mechanism of trust.

We turn in the next section, however, to a mechanism that does not

10. Avner Greif, “Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders,” *Journal of Economic History* 49 (1989): 857–82.

11. *Ibid.*, p. 862.

12. *Ibid.*, p. 867.

13. *Ibid.*

depend on a belief in the trustworthiness of the trustee. It involves a belief that the trustee will be reliable under trust, of course, but the reliability posited does not spring from the possession of what is generally taken as a desirable trait and does not count in our sense as a form of trustworthiness. The mechanism, as we shall see, is parasitic on the mechanisms we have just described. But it is a mechanism of perhaps even greater importance, as should become clear in later discussions. It can work in the service of trust, even when traits like loyalty, virtue, and prudence are in short supply or under severe pressure.

III. A MECHANISM OF TRUST-RESPONSIVENESS: REGARD-SEEKING

There are two fundamentally different sorts of goods that human beings seek for themselves. The one kind may be described as attitude-dependent, the other as action-dependent.¹⁴ Attitude-dependent goods are those which a person can enjoy only so far as they are the object of certain attitudes, in particular certain positive attitudes, on the part of others, or indeed themselves. They are goods like being loved, being liked, being acknowledged, being respected, being admired, and so on. Action-dependent goods are those which a person can procure without having to rely on the presence of any particular attitudes in themselves or others; they are attained by their own efforts, or the efforts of others, and they are attained regardless of the attitudes at the origin of those efforts. Action-dependent goods are illustrated by the regular sorts of services and commodities and resources to which economists give center stage.

Although economics focuses on action-dependent goods, it should be clear that people care also about goods in the attitude-dependent category; they care about being cherished by others, for example, and about being well regarded by them.¹⁵ It is striking, indeed, that Adam Smith, the founding father of economics, seems to have thought that the desire for the good opinion of others, the desire for standing in the eyes of others, was one of the most basic of human inclinations. "Nature, when she formed man for society, endowed him with an original desire

14. Pettit, *The Common Mind*, Chap. 5.

15. Pettit, "Virtus Normativa Rational Choice Perspectives," *Ethics* 101 (1990): 725–55; Geoffrey Brennan and Philip Pettit, "Hands Invisible and Intangible," *Synthese* 94 (1993): 191–225.

to please, and an original aversion to offend his brethren. She taught him to feel pleasure in their favorable, and pain in their unfavorable regard. She rendered their approbation most flattering and most agreeable to him for its own sake; and their disapprobation most mortifying and most offensive.”¹⁶ Smith even seems to have held that the reason people seek more and more goods in the action-dependent category is that such goods serve to confer distinction and standing.

I am going to assume that Smith is right in thinking that people do seek the good opinion of others, even if this desire is not any more basic than their desire for material goods. My view is that desires for attitude-dependent goods and desires for action-dependent goods are probably of equal status: furthering each sort of desire is satisfying in itself, so that each can be seen as a basic desire; and furthering each probably represents an indirect way of furthering the other, so that each can also be seen as having derived or instrumental significance. But nothing depends in what follows on that view. All I need to assume here is that people do have desires for attitude-dependent goods, in particular desires for the good opinion of others. I can remain uncommitted on whether that desire is basic or on whether its strength depends on the fact that by getting others to think well of them, people are better able to secure the material goods they pursue.

Let us assume, then, that each of us desires the good opinion of others. The availability of a further mechanism of trust becomes visible as soon as we ask whether that assumption might give a trustor independent reason to expect that a trustee will be motivated by the trustor’s manifestation of relying on the trustee: in particular, reason to expect that a trustee will be motivated to perform as the trustor relies on her to perform. For the answer to that question is clearly that sometimes there is reason, related to the desire of a good opinion, why a trustor might expect her manifestation of reliance to be motivating: specifically, to raise the utility that the trustee attaches to proving reliable.¹⁷

In some circumstances the manifestation of reliance may communicate a belief, not that the person relied on is trustworthy, but only that he is bound by such constraints that he will behave as required; this is

16. Adam Smith, *The Theory of the Moral Sentiments*, ed. D. D. Raphael and A. L. Macfie (Indianapolis: Liberty Classics, 1982), p. 116.

17. Anthony Pagden, “The Destruction of Trust and its Economics Consequences in the Case of Eighteenth Century Naples,” in Gambetta, ed., *Trust*, p. 133.

liable to happen in the bus-driver case, for example, when the driver is independently obliged to go to the city center. But the manifestation of trusting reliance, the manifestation of reliance which also manifests a belief that the person relied on will be motivated by the fact of reliance to prove reliable, may well be different. It can be a token offered by the trustor of believing the trustee to be trustworthy, or of being disposed to believe this in the event of the trustee's proving reliable: it can communicate a judgment that the trustee is trustworthy or at least a presumption, as we may call it, on such trustworthiness.

The trustor will not typically utter words to the effect that the trustee is someone who will not let the needy down: that the trustee, as we say, is indeed a trustworthy individual. But what the trustor does in manifesting trusting reliance may be tantamount to saying something of that sort. Let the context be one where the trustor can be taken to expect the trustee to prove reliable only if the trustee has a modicum of trustworthiness: only if the trustee is or proves to be loyal or virtuous or prudent. In such a context the act of trust will be a way of saying that the trustee is indeed a trustworthy sort.

Indeed it will be something of even greater communicative significance, for words are cheap and actions dear. The act of trust will communicate in the most credible currency available to human beings—in the gold currency of action, not the paper money of words—that the trustor believes the trustee to be truly trustworthy, or is prepared to act on the presumption that he is: believes or presumes him to be truly the sort of person who will not take advantage of someone who puts herself at his mercy. It does not just record the reality of that attitude, it shows that the attitude exists.

To think that someone is trustworthy, whether in the way of loyalty or virtue or prudence, is ordinarily to think well of him; it is to hold him in high regard. Traits like loyalty and virtue and prudence are, by all accounts, desirable traits, at least when they are given their proper place.¹⁸ Thus the fact that I manifest trusting reliance in an act of trust—if I do so—means that as I communicate a judgment or a presumption

18. There are some complexities here. I may think you loyal enough to do me a favor in your capacity as a public official. But to think you trustworthy in that way may not be, by our shared lights, to think well of you; after all, the loyalty ascribed is misplaced loyalty. I abstract from this complexity in my discussion. It has significance for the dynamics of trust among agents who are consciously in violation of certain norms: consciously non-virtuous in a certain manner.

of trustworthiness, I can communicate that I think well of the trustee or at least that I will do so in the event of his proving reliable. When it connects in this way with the desire of a good opinion, then the act of trust is likely to have an important motivating aspect for the trustee.

In such a case the act of trust makes clear to the trustee that he enjoys or will enjoy the good opinion of the trustor—the belief that they are trustworthy—just so long as he behaves in the manner required. This means that the trustor has a reason to expect the manifestation of reliance to be motivating with the trustee, independently of any belief in his pre-existing loyalty or virtue or prudence. If the trustee values the good opinion of the trustor, then that is likely to give them pause about letting the trustor down, even if he is actually not a particularly loyal or virtuous or prudent person. Let the trustor down and he may gain some immediate advantage or save himself some immediate cost. But let the trustor down and he will forfeit the benefit of being well regarded by the trustor: that, and all the other benefits that may be associated with sustaining such a good opinion.

But there is also more to say. By displaying trust in another, one often demonstrates to third parties that she trusts that person. Other things being equal, such a demonstration will serve to win a good opinion for the trustee among those parties; the demonstration will amount to testimony that the trustee is a trustworthy person or is worthy of being given the chance to prove himself trustworthy. Indeed if the fact of such universal testimony is salient to all, the demonstration may not only cause everyone to think well of the trustee; it may also cause this to become a matter of common knowledge, thereby giving the trustee the public status of a trustworthy person. Assuming that such facts are going to be visible to any perceptive trustee then, the existence of independent witnesses to the act of trust will provide further regard-centered motives for them to perform as expected. Let the trustor down and not only will they lose the good opinion that the trustor has displayed or promised; they will also lose the good opinion and the high status that the trustor may have won for them among third parties.

The other mechanisms of trust all explain why any risk-taking that trust involves may actually be quite sensible. Maybe there is a risk involved in this or that act of trust but the risk is not substantial—it is, at the least, a rational gamble—given that the trustee is suitably loyal or virtuous or prudent. While the present mechanism also explains why

such risk-taking may be quite sensible, it does so in a distinctive manner. To manifest trusting reliance, so it now appears, is to provide the trustee with an incentive to do the very thing that the trustor is relying on them to do. It is a sort of bootstraps operation, wherein the trustor takes a risk and, by the very fact of taking that risk, shifts the odds in their own favor.

As Hegel spoke of the cunning of reason, so we can speak here of the cunning of trust. The act of trust is an investment by the trustor which will pay dividends only in the event that the trustee behaves appropriately. Like any investment it may have a risky side, for the trustee may not be bound to act as required. But it is not by any means as risky as it may at first seem. For in the very act whereby the trustor is put at risk, the trustee is given a motive not to let that risk materialize. The trustor can bank on the fact that if the trustee does let the risk materialize then they will suffer the loss of the trustor's good opinion and, in all likelihood, the cost of gaining a bad reputation among those who learn of what has happened.

It may be useful, in summary, to offer a brief, premise-by-premise statement of the argument.

1. There are situations where an act of trust will signal to a trustee, and to witnesses, that the trustor believes in or presumes on the trustworthiness of the trustee—believes in or presumes on his loyalty or virtue or prudence—and so thinks well of him to that extent.
2. The trustee is likely to have a desire, intrinsic or instrumental, for the good opinion of the trustor and of witnesses to the act of trust.
3. The desire for that good opinion will tend to give the trustee reason to act in the way in which the trustor relies on him to act.

Conclusion. And so the trustor, recognizing these facts, may have a reason to trust someone, even when he actually has no reason to believe in the other's pre-existing trustworthiness.

Where the mechanisms described in the last section were described as mechanisms of trustworthiness, the mechanism to which our attention has been directed here is one of trust-responsiveness. The reason the trustor believes that his manifesting reliance will motivate the trustee is that it is a manifestation of trusting reliance which communicates a belief in, or a presumption on, the trustworthiness of the trustee. The trustor thinks, not necessarily that the trustee has one of the desirable

traits that make for trustworthiness, but rather that the trustee will be affected by the act of trust and that this will give him a reason to prove reliable; the trustor thinks, in a word, that the trustee is responsive to acts of trust, specifically to acts of trust that manifest trusting reliance, and that this is a good ground for placing trust in him.¹⁹

We mentioned earlier that the mechanisms of trustworthiness can be mutually reinforcing and should not be construed as in competition. A similar point may be made about the mechanism of trust-responsiveness in relation to those other three. There is no difficulty in the idea that a trustor might take a trustee to be simultaneously moved by trustworthiness and by trust-responsiveness: to be moved at once by loyalty and by the desire to be thought loyal, by virtue and by the desire to be thought virtuous, by prudence and by the desire to be thought prudent. And equally there is no difficulty in the idea that a trustor may take a trustee to be moved by a trustworthiness trait, while deriving confidence from the thought that if that trait fails to operate spontaneously—if loyalty or virtue or prudence slips—still the trust-responsiveness mechanism is waiting in the wings to help out: as the person becomes aware of the loss of standing that a betrayal of trust will entail, the desire for such standing may cut in and ensure a suitable performance.²⁰

IV. THE EXPLANATORY POTENTIAL OF TRUST-RESPONSIVENESS

The three trustworthiness mechanisms pose a number of explanatory questions with which the trust-responsiveness mechanism helps us to deal. The questions bear respectively on the phenomenology, the ubiquity, and the creativity of trust. To take the phenomenology first, the standard, trustworthiness mechanisms let us see how acts of trust, in particular acts of trust that are not betrayed, can be epistemologically self-reinforcing, offering the trustor and the trustee increasing reason to

19. Is trust-responsiveness likely to be underpinned only by the desire for the good opinion of others? Perhaps not. The desire of your own good opinion may also play a role. People may do things, not just for the sake of winning or keeping the good opinion of others, but also for the sake of being able to think well of themselves. Thus you may think that by manifestly trusting another you can put that person in a position where, if they let you down, they have to think badly of themselves. My thanks to Dan Hausman for a discussion of this point.

20. Pettit, *The Common Mind*, Chap. 5; idem, "The Virtual Reality of Homo Economicus."

believe that neither will let the other down. But they fail to explain the fact that trust also appears to be motivationally self-enforcing: It does not serve just to give trustor and trustee greater confidence in one another—specifically, in one another's loyalty or virtue or prudence—it also has the effect of binding them more closely with each another. It is a pleasure to find oneself trusted by someone, and it is a pleasure for others to find themselves trusted by you, and none of the trustworthiness mechanisms suggests why this should be so.

The standard mechanisms of trust also fail to make good sense of the ubiquity of trust in civil society. We do not trust only those of whom we have prior reason to believe that they are loyal or virtuous or even prudent. Nor do we exercise trust just in situations where our general knowledge of the culture or the institutions or the people gives us indirect reason to believe that certain individuals are trustworthy. We also trust people with whom we may have had little to do, and of whom we may have little direct or indirect knowledge. Think of the new resident who asks a neighbor to look after her home or pets or plants while she is away and gives them a key to her house. Think of the passenger who admits to not knowing the town and asks a taxi driver to get him to his destination by the quickest route available. Think of the person who asks a perfect stranger for directions on how to get somewhere and then follows these meticulously. Think of the customer who, finding that one store does not have something he wants, asks the salesperson for advice on where else to search. Or think of the visitor who asks a newsagent to recommend a good evening's read or asks a cinema attendant's opinion of the film showing. Such expressions of trust, however trivial, are characteristic of flourishing civil societies.²¹ But it is hard to see how they can be fully explained by mechanisms which presuppose that the trustor already has reason to believe in the trustworthiness of the trustee.

The standard mechanisms fail, finally, to explain the creative aspect of trust. Not only does trust build on trust, as in the accumulation of trust within a given relationship, trust can also build on nothing and can help to establish such relationships in the first place. It can create *de novo*. Think of the trust expressed in an overture of friendship, as when someone asks another person for very personal advice and relies on the other, not just to offer his best counsel, but also to respect his confi-

21. Robert D. Putnam, *Making Democracy Work: Civic Traditions in Modern Italy* (Princeton: Princeton University Press, 1993).

dence. Think of the trust expressed in the approach to someone which treats them as a worthy and reliable individual, though the trustor does not have independent knowledge that the trustee really is possessed of such virtue. Or think of the trust expressed in the business gamble that supposes that another person—a person whom the trustor scarcely knows—will believe that there are long-term benefits in the offing, and will make the gamble come good.

The phenomenology, the ubiquity, and the creativity of trust become less surprising in the light of trust-responsiveness. Why should people take pleasure in being trusted: specifically, take a sort of pleasure that seems distinct from the epistemic pleasure of becoming more confident that the trustor can be trusted in turn? If trusting someone is a way of communicating a good opinion of her, and if people savor the good opinion of others, then there is no difficulty in seeing why. People take pleasure in being trusted, because people take pleasure in being well considered and well regarded.

How can trust occur outside the realms where people have direct or indirect reasons for believing in the loyalty or virtue or prudence of those in whom they invest trust? Even where people have no independent reasons for positing trustworthiness, they may have reason to assume a desire for regard: this, to the extent that such a desire is a robust feature of human psychology. And if they assume such a desire, they will often have reason to expect acts of trust to motivate trustees. Thus the domain where they are prepared to exercise trust may considerably out-run the domain where the trustees are proven figures of loyalty or virtue or prudence.

As against this suggestion, it may be said that I am unlikely to care for the regard that a perfect stranger communicates by an act of trust and, seeing that it is given so lightly, may even dismiss it as insignificant. But a number of points are worth making here. One is that the desire for regard seems to operate quite robustly with strangers—think of the embarrassment of being seen picking your nose by someone you don't know—so that I may shrink from letting even a total stranger down and thereby forfeiting their good opinion. Another point is that I may well feel that the stranger who trusts me likes the look of my face, so that there is a nontrivial, personal basis for the regard communicated. And a third is that I may think that the stranger trusts me on the grounds of my group affiliation—I am a clergyman, I am black, I am a hippy, I am

a local—and that in such a case I will have a reason of collective identification to care for sustaining the stranger's good opinion.²²

How, finally, can trust be creative, helping to establish relationships in which there is a common belief among the parties involved in their loyalty or virtue or prudence? A relationship of mutually acknowledged loyalty, or a relationship involving mutually recognized virtue or prudence, can get established just to the extent that one party can credibly communicate a belief in the loyalty or virtue or prudence of the other, and can set in train a process whereby that belief is reinforced on both sides and comes to be shared in common between them. That communication may be effected in words. But it can also be effected, and effected without irrational risk, in actions of trust. For as I invest another with a certain trust, I can communicate the sort of belief that may lead to a corresponding relationship being established. And I can rationally invest that trust in advance of the relationship being formed, and without knowing whether the other is trustworthy, given that the act of trust can prove inherently motivating: can provide an incentive in the economy of regard for the trustee not to let me down.

V. THE INSTITUTIONAL SIGNIFICANCE OF TRUST-RESPONSIVENESS

Our analysis of trust, and in particular of the trust-responsiveness mechanism whereby trust may be generated and sustained, has implications for institutional design. Consider the three premises in the argument for the trust-responsiveness mechanism, as that argument was summarized at the end of Section III. Each of these premises requires that certain conditions obtain if it is to hold good. And whether those conditions hold good is often a function of how institutional matters are designed and arranged.

The third premise holds that the desire for the good opinion of the trustor, and of witnesses to the act of trust, will tend to give the trustee reason to act in the way in which the trustor relies on him to act. Yes, but only if things are arranged in such a way that it is obvious whether the trustee does indeed behave in the required manner. And arranging things in that way may mean ensuring that suitable analysis and information of the trustee's performance is made available to trustors. This

22. Thanks to Peter Godfrey-Smith for this point.

can be of great institutional significance if we are trying to devise institutions under which people can be confident about investing trust in commercial organizations, medical advisors, environmental agencies, and the like.

The second premise in the argument for the trust-responsiveness mechanism holds that the trustee is likely to desire the good opinion of the trustor and of witnesses to the act of trust. Yes, but only if there is not a division in the community, in particular a division between the trustee and those others, which makes people on one side indifferent to what people on the other think of them. It is all too obvious that divisions of creed and color and gender, and any of a myriad of political divisions, can undermine community between people to such an extent that neither side cares about being thought to behave shamefully by the other. Among the many ills that such a division can bring about, it is liable to reduce the chances of trust materializing between people from the different sides. It is liable to inhibit the operation of the trust-responsiveness mechanism, as those on each side become indifferent to the good opinion of those on the other, and as this indifference comes to be a matter of common recognition.

But the institutional lessons underpinned by a recognition of the trust-responsiveness mechanism become most telling, as we look at the conditions necessary for the first premise in our argument to hold good. This premise says that an act of trust can signal to a trustee, and to witnesses, that the trustor believes in or presumes on the trustworthiness of the trustee and thinks well of her to that extent. Three conditions are clearly necessary if an act of trust is to fulfill this signalling function, and it is important for anyone concerned with institutional policy to be aware of these; otherwise they may advance policies that would undermine trust, or may fail to advance policies that would encourage it.

A first condition necessary for an act of trust to communicate a belief or presumption that the trustee is trustworthy is that there be enough instances of trustworthiness in evidence, and enough knowledge of those instances, for it to be plausible that someone should believe that another person is trustworthy. Let trustworthiness be a very scarce resource, or let it be thought to be a very scarce resource, and it is not going to be plausible that a person will take himself to be regarded as trustworthy just because another manifests a certain reliance on him. To the extent that a community fails to display examples of loyalty and

virtue and prudence, then, and of relationships organized around them, it is less likely to have much of a place for trust. Not only will the failures in these traits, and in the relationships that the traits support, mean that there are correspondingly fewer outlets for the standard mechanisms of trust. They will also mean that the prospects for overtures of trust that are supported by the mechanism of regard-seeking are dramatically reduced. In a society where there are fewer examples of trustworthiness—fewer examples of relationships and institutions built around attributions of loyalty or virtue or prudence—there will be weaker inclinations on the part of trustees to think that they are regarded as trustworthy or on the part of trustors to expect that trustees will think this.

To them that have, it shall be given. Where there are already lots of examples of trust and of trusting relationships in a society, there will be correspondingly greater opportunities for people to exploit one another's desire for regard and to let trust innovate and develop. Where a society has degenerated to the point that there are few institutions of trust, it is hard to see how things may be transformed so as to let trust in. Consider the society, for example, where trust is only found in small family groups: where there are few other examples of loyalty-based trust and few or no examples of trust based on habits of expecting virtue or prudence.²³ Consider a society, in other words, where civic engagement is at an absolute minimum and utter cynicism prevails: where there is little of what James Coleman²⁴ describes as social capital. In such a society, trust is likely to lack any dynamic and it may require dramatic developments or interventions if things are to be turned around.

One way in which a society might become utterly cynical and might undermine the trust-responsiveness mechanism offers us a nice paradox. Suppose that it became a matter of common belief in the society that no one was trustworthy and that the only reason anyone trusted anyone else was the belief that this would communicate a good opinion of the trustee and exploit the trustee's desire to secure that good opinion by proving reliable. In such a case, paradoxically, people would cease to think that being trusted was a case of being well regarded. For whereas it may be a complement to be thought loyal or virtuous or prudent, it is no complement to be thought to want the good opinion of others.

23. Gambetta, "Mafia: The Price of Distrust," in Gambetta, ed., *Trust*.

24. James Coleman, *The Foundations of Social Theory* (Cambridge, Mass.: Harvard University Press, 1990), pp. 300–21.

“The general axiom in this domain is that nothing is so unimpressive as behavior designed to impress.”²⁵

A second condition necessary for the trustor’s manifestation of reliance to communicate a belief or presumption that the trustee is trustworthy is that the trustor does not have any more salient motives for manifesting reliance. Suppose that the trustor is a subordinate who is utterly at the mercy of the trustee. Any attempt by such a trustor to communicate a belief or presumption that the trustee is trustworthy is liable to be seen as a fawning act, designed to placate the trustee. Or suppose that the situation is the reverse, so that the trustee is the subordinate and lives at the mercy of the trustor. Any attempt by such a trustor to communicate a belief or presumption that the trustee is trustworthy is liable to be seen as a sort of condescension, designed to make the trustee feel good. In circumstances of either kind the trustor will be unable to communicate a good opinion of the trustee. There will be too much noise in the channel.

The lesson of this condition is that it will be difficult for anyone to manifest trusting reliance on another, and thereby motivate that other to prove reliable, if they are utterly vulnerable to that person or if that person is utterly vulnerable to them. From the point of view of the stronger, the weaker’s apparently flattering act of trust is easily seen as a sort of sycophancy or self-ingratiation. The weaker person has need of the good will of the other and the act of trust is easily put down to a self-abasing attempt to win favor, in which case it will fail to be the motivator that it can be when practiced between equals. From the point of view of the weaker, on the other hand, the stronger person’s apparently flattering act of trust is likely to fail in a complementary way. The weaker person will reckon that within suitable limits the stronger must expect them to satisfy the stronger’s wishes, however those wishes are communicated, and will see the alleged act of trust as a more or less hypocritical routine: an indirect way of commanding the response sought which only a simpleton could take as the compelling expression of a good opinion.

If we want to maximize trust then we should look for a society, large-scale or small-scale, in which no one is forced to live at the mercy of others, a microcosm or macrocosm in which people enjoy freedom as

25. Jon Elster, *Sour Grapes* (Cambridge: Cambridge University Press, 1983), p. 66.

non-domination.²⁶ Only by guarding against enforced vulnerability can a society facilitate the voluntary assumption of vulnerability—the voluntary assumption of limited vulnerability—that is associated with trust. Only by being buried, can the seed bring forth life.

A third condition necessary for the manifestation of reliance to communicate a belief or a presumption that the trustee is trustworthy is that the trustee is not subject to such pressures to act in the required way that any manifestation of reliance is more plausibly explained as stemming from a recognition of those pressures. The situation must not be like the one where the bus driver on whom I rely—the bus driver whom I may actually trust—to take me to the city center is independently constrained to go there. Let the trustee be coerced or constrained to do something, A, and it is going to be very difficult for someone to communicate a belief in his trustworthiness just by making it clear that he relies on him to do A. Thus it is going to be difficult or impossible for a person to exercise trust on the basis of the trust-responsiveness mechanism.

Imagine a circumstance where someone can manifest trusting reliance on another to act in a certain way. And suppose now that big sanctions are put in place which make it very likely, independently of any trustworthiness, that the person relied upon will comply with expectations. Suppose things are rigged by the sanctions, in other words, so that it would be quite irrational for the person relied upon not to satisfy the other. Will the trustor continue to be able to communicate a belief in the trustworthiness of the other person by manifesting reliance on him? Surely not. The more likely explanation of the manifestation of reliance in such a case will always be that the trustor expects the trustee to be motivated by the sanctions: that is, expects the trustee to behave in the rational, self-interested way.

The point is readily illustrated. Imagine the difference that can be made when an organization introduces various checks on when their nonmanagerial staff turn up for work and how they spend their time. Previously a manager in such an organization might have expressed trust in one of their staff by giving her some task to perform that would allow her, if she so wished, to exploit the trustor: to take an excessive amount of time over the job, to do the job sloppily, or whatever. Previously the expression of such trust, flattering as it is, might well have led

26. Pettit, "Freedom as Antipower," *Ethics* 106 (1996).

to a relationship of trust between the manager and the member of staff, with all the attendant benefits that that can bring. But now that the checks have been put in place, the opportunity for the manager to manifest trusting reliance in the member of staff has been removed. The checks mean that the member of staff will have salient and unflattering reasons to comply, so that the manager's request cannot have the aspect of an expression of trust and cannot serve to establish a trusting relationship between the two.

This final lesson is important, because it shows how certain intrusive forms of regulation can be counter-productive and can reduce the level of performance in the very area that they are supposed to affect. The ways in which heavy regulation may be counterproductive are various,²⁷ but I suspect that this is one of the most important. If heavy regulation is capable of eradicating overtures of trust, and of driving out opportunities for trusting relationships, then it is capable of doing great harm.

We have just been looking at conditions that are necessary for the manifestation of reliance to communicate a belief or presumption that the trustee is trustworthy. The first was that there are enough instances of trustworthiness in evidence to make it plausible that a trustor should hold by such a belief or presumption. The second condition was that the trustor does not have any more salient motives for manifesting reliance. And the third condition was that the trustee is not subject to such pressures to act in the required way that any manifestation of reliance is more plausibly explained as stemming from a recognition of those pressures. These conditions, and the others mentioned earlier, represent different requirements for the smooth functioning of the trust-responsiveness mechanism in generating and sustaining trust. They point up some important lessons for institutional designers, since it is now common wisdom that trust is a precious if fragile commodity in social and political life.²⁸ Institutional policy-makers and designers ignore such lessons at their peril.

27. Ian Ayres and John Braithwaite, *Responsive Regulation* (New York: Oxford University Press, 1992); Peter N. Grabosky, "Counterproductive Regulation," *International Journal of the Sociology of Law* 23 (1995); and Cass R. Sunstein, "Paradoxes of the Regulatory State," *University of Chicago Law Review* 57 (1990): 407-41.

28. Dasgupta, "Trust as a Commodity," in Gambetta, ed., *Trust*.