Trustworthiness

Russell Hardin


Stable URL:
http://links.jstor.org/sici?sici=0014-1704%28199610%29107%3A1%3C26%3AT%3E2.0.CO%3B2-9

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. 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.

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.

*Ethics* is published by The University of Chicago Press. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ucpress.html.

*Ethics*
©1996 The University of Chicago Press

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2002 JSTOR
Trustworthiness*

Russell Hardin

THE CURSE OF ALBERICH

The cataclysmic events of Wagner's cycle of operas, *The Ring of the Nibelung*, follow from an extraordinary act of commitment in the opening minutes of *The Rhinegold*. Alberich, the Nibelung, feverish with lust for the Rhine maidens, learns that anyone who can forswear love forever can gain enormous power from the Rhinegold. Instantly he seizes the gold, swearing, "Love, I curse you forever!" By committing himself forever, Alberich gains extraordinary power. Even the great god Wotan, who is unable to make such a commitment, is impressed by the lowly Alberich's resolve.²

There is, of course, much in Wagner that strains belief. But is anything in the entire *Ring* less credible than the possibility that any young adult male could turn from lusting to the point of rapacious madness in one minute to credibly forswearing lust forevermore in the next? Alberich, one of the pettiest figures in all of literature and opera, makes one of the grandest of commitments. Could anyone trust Alberich's conversion? If we are reasonable, we should trust him only if we believe his commitment will carry him through. But could one reasonably trust his commitment? Unfortunately, Wagner does not tell us enough to know. For all its greatness, the *Ring* turns on this bit of psychological nonsense.

Alberich is preposterous. If we wish to understand real people, we will need a better psychology. What we will have to understand are the capacities for making commitments and trusting those who make them, capacities which must largely be learned. One cannot simply

---

*This article was presented to a conference on trust at the Institute of Cognitive and Decision Sciences, University of Oregon, November 1994. I thank the organizers of that conference for the occasion and the participants for comments on the paper.

1. Richard Wagner, *The Ring of the Nibelung*, trans. Andrew Porter (New York: Norton, 1977), p. 16. Many of Wagner's characters, some of them among the most loathsome and treacherous in all of opera, are obsessed with trust. Like Lohengrin, they demand it without evidence that they merit it; like Wotan, they are unworthy of it.

2. Ibid., pp. 106–7, in *The Valkyrie*.

*Ethics* 107 (October 1996): 26–42
© 1996 by The University of Chicago. All rights reserved. 0014-1704/97/0701-0002$01.00
start trusting people as of tomorrow unless the people one deals with are suddenly different in relevant ways. When I meet someone new with whom I wish or have to deal, I may start with considerable skepticism. But my skepticism will not be directed primarily at the new person in particular. I may not yet know enough about her to judge her trustworthiness or her rationality in being trustworthy. I make my skeptical judgment largely by generalization from past encounters with other people. In that sense, my degree of trust in the new person has been learned.

My prior experiences with trust may have been so charmed that I optimistically trust this new person. Or they may have been so disastrous that I pessimistically distrust her. She is no different in the two cases; my alternative experiences are the source of difference. Experience molds the psychology of trust. If my past experience too heavily represented good grounds for trust or distrust, it may now take a long run of contrary experience to correct my assessments and, therefore, my actual psychological capacities.

What can that person do to convince me of her commitment to fulfilling my trust? For example, it might have been in Alberich's interest to say he would forswear love forever even if he knew he would later cheat. What could have made his curse credible? If he genuinely forswore love forever, imagine how forlorn he must have been when, four operas later, he was one of the few survivors in the cast to see the delectable Rhine maidens retrieve their gold. Poor Alberich, now bereft of the gold and its power, was still bereft of the possibility of love. Perhaps it is plausible that some persons could actually make as dramatic a commitment as Alberich made and then stick with it merely from strength of character or some odd capacity for consistency. David Gauthier seems to think we can will ourselves to have a particular disposition that, in the moment we bring it to bear, is not in our interest. But most of us would probably have to admit, if we are honest, that we cannot do such a thing. We might think ourselves capable of undying love for and commitment to, say, our child or parent. But we do not come by that love through mere decision, as Alberich supposedly came by his commitment.

Few of our commitments ever have such force as Alberich's curse supposedly had. Our strongest commitments are often merely those that are clearly backed by our interest. For example, Alberich might have been governed by some magical power that would have taken back the Rhinegold if he ever acted on his lust. Then, at least, perhaps he ended the Ring cycle without the Rhinegold but with the renewed

possibility of love. Our task is to contrive real-world substitutes for such magical powers as Alberich faced to keep us in line.

TRUST AND TRUSTWORTHINESS

Many discussions of trust run trust and trustworthiness together, with claims about trust that might well apply to trustworthiness but that seem off the mark for trust. In particular, writers often transfer to trust the moral approbation that might be thought applicable to trustworthiness. Independently of whether there is something moral about being trustworthy or untrustworthy, however, trust might be fully explicable as a capability or as a product of rational expectation without any moral residue. I treat trust as an unmoralized notion, as Coleman and many others commonly do. Most of my argument would have some force even in a moralized account of trust in which the potential truster is held morally accountable for a failure to trust. My argument would especially be of interest in such a moralized account insofar as the argument is about constraints on the capacity to assess trustworthiness.

Surprisingly, much of the literature on trust hardly mentions trustworthiness even though much of it is primarily about trustworthiness, not about trust. Consider three striking cases. First, under the guise of discussing trust, the philosopher Bernard Williams gives an account of the possibilities of general trustworthiness, from which trust is merely inferred. Second, the economist Roland McKean ostensibly addresses the economics of trust, but his actual problem is that of trustworthiness. It is not trust per se that is the collective good in his account, but trustworthiness. Creating institutions that help secure trustworthiness thus helps to support or induce trust.

The third case is more complicated. The sociologist Niklas Luhmann says trust constitutes a more effective form of complexity reduction. This seems to be a very elliptical claim. Clearly, where there is trust that is justified there are increased possibilities for beneficial experience and action. Trust by itself, however, constitutes nothing. Presumably, Luhmann is saying that we cannot handle enormous complexity without having others act de facto on our behalf. But if we cannot count on their acting in our interest, we may therefore be reluctant to empower them or to follow their advice. Hence, whatever

can secure their trustworthiness enough for us to trust them will help us manage complexity. That is to say, again, the focal problem is trustworthiness, not trust.\textsuperscript{7} Elsewhere, Luhmann says he is concerned with social mechanisms that generate trust.\textsuperscript{8} Once again, he has substituted concern with trust for concern with trustworthiness. The best device for creating trust is to establish and support trustworthiness. As before, without the latter, there is no value in trust.

Linger with this issue for a moment. In \textit{The Remains of the Day} Kazuo Ishiguro describes Mr. Stevens, an aging butler rethinking his life with his late master. In an imagined debate with another servant Stevens says, “The likes of you and I will never be in a position to comprehend the great affairs of today’s world, and our best course will always be to put our trust in an employer we judge to be wise and honorable, and to devote our energies to the task of serving him to the best of our ability.”\textsuperscript{9} He slowly revalues his master in the light of the views of others who detested the lord’s reprehensible and foolish politics. Stevens says, “At least [his lordship] had the privilege of being able to say at the end of his life that he made his own mistakes . . . I cannot even claim that. You see, I trusted. I trusted in his lordship’s wisdom. All those years I served him, I trusted I was doing something worthwhile. I can’t even say I made my own mistakes. Really—one has to ask oneself—what dignity is there in that?”\textsuperscript{10}

Stevens recognizes that trust can finally be stupid and, when it seemingly justifies action or inaction, even culpable. Merely trusting per se obviously need not help in managing complexity well—it could lead to dismal results, including quick destruction. Again, the core of Luhmann’s account of the role of trust must be an account of the importance of trustworthiness. If his account really commends trust in its own right, it elevates Stevens’s culpable stupidity. Trust led Stevens not to manage complexity so much as to fall victim to it.

\textbf{COMMITMENT}

The central problems of effective cooperation are to commit oneself and to convince others that one is committed. Capacity for making and communicating a commitment is of much more general interest

\textsuperscript{7} Luhmann also says that trust is a kind of capital (ibid., p. 64). While trustworthiness might be seen as a form of capital, as reputation is, it is hard even to imagine what it means for trust to be a form of capital. That we all trust might seem to yield good results, so that we might think generalized trusting is a form of capital. But, again, good results depend on trustworthiness, without which generalized trusting cannot and should not arise.

\textsuperscript{8} Niklas Luhmann, “Familiarity, Confidence, Trust: Problems and Alternatives,” in Gambetta, pp. 94–107, p. 95.


\textsuperscript{10} Ibid., p. 243.
than merely for establishing trustworthiness, and we can approach the more specific problem from the more general. Alberich’s curse could be seen as a commitment to trustworthiness. Or it could merely be seen as a commitment not to act from lust, a commitment to abjure love. Keeping our focus on the latter makes it more natural to see just why the commitment is dubious. Alberich’s lust is largely not a matter of deliberate choice. Rather, it is something at least partly in control of him, something that gives him strong motivations. His commitment to defy it means a commitment to ignore those motivations when they come. Typically, we think such a strong motivation cannot be literally ignored, although it can be trumped by a contrary motivation, such as Alberich’s desire for wealth and power. His psychological problem, once he has the Rhinegold, is to connect future motivations of lust to his contrary desire for power, so that his lust can be trumped.

Either there is a magic force external to Alberich that will cancel his wealth and power when he slips and acts on his lust, or his commitment is entirely internal. Insofar as we can know what Wagner might have meant about something that he did not write in sufficient detail to tell us directly, we can suppose that Alberich’s commitment was strictly internal. He willed himself not to entertain love ever again. Hobbes and many others ridicule such willing. As Hobbes noted, there is nothing that keeps me from changing my mind if there is only my own will to motivate me. If I am sovereign, I can be capricious. This seems so overwhelmingly correct a view that there is little point in exploring the possibility of direct willing of trustworthiness.

Nevertheless, one can sometimes “will” oneself into future action indirectly. For example, one can predict well enough how one would act in certain circumstances and can therefore sometimes set matters up to stimulate a particular response. For example, I may know that, once I get out of bed in the morning, I am really up and I will not return to sleep. Therefore, I set my alarm clock on the dresser a few feet from my bed where I will have to stand to stop its noise. I have then used an external device to get myself on my feet, but it is something internal about me that keeps me from returning to bed and to sleep. If the alarm were too near, I might often turn it off and go back to sleep. You might be even better than I am at getting up, and you might find yourself uncomfortable to continue in bed after first waking, so that the nearby alarm would work for you. Perhaps analogous devices can be called into action for making ourselves trustworthy in certain matters. For example, I might make my commitment in public in full self-understanding that my embarrassment over publicly failing the commitment would trump my costs of keeping to it. Still, the range of such devices is apt to be relatively limited, so that I am not likely to be able to use internal devices other than pure will to manipulate myself into being trustworthy much of the time.
Henceforth, let us suppose that the devices that secure our important commitments are largely external except for the role our current motivations play in our current actions. We therefore are concerned with social constraints. Society offers two general or modal categories of controls to individuals who can benefit from constraining themselves. There are the particular, small-scale controls of ongoing relationships of family, friends, and what we might crudely call geographical associates—those with whom we will almost inescapably be thrown into further dealings. And there are the elaborate, large-scale controls of the law and other institutions in a relatively extensive society. Between these two modal categories there are mixed devices such as institutionalized religious controls and broad social norms and practices that blend elements of both the modal categories. Where these social controls fail we are left with our own personal devices of internal motivations, which all too often means we are left in the lurch.

The first of these social devices—constraint by our ongoing relations with close associates—is spelled out elsewhere.\textsuperscript{11} It involves the reinforcing incentives that come from ongoing relations, such as iterated exchange relations. In such a relationship, my reliability in this moment may be reinforced by my interest in having the relationship continue. The second device—constraint by constructed organizational incentives—is discussed immediately below. The third, mixed device—constraint by larger social conventions—is discussed immediately thereafter. The focus here is on the use of whatever devices, including organizations, can be made to support trust between persons by making them trustworthy. For example, the legal system of contract enforcement enables me to trust you in some formal exchange. This raises the questions whether I can trust a particular institution and whether the institution has capacity for trustworthiness, but I will not discuss this complex problem here.

INSTITUTIONAL CONSTRAINTS

If I trust you to act on my behalf, I set myself up for the possibility of disappointment, even severe loss. To avoid that possibility, I might try to find institutional backing to get you to do what I trust you to do. This may be very easy to do. For example, if I wish to trust you to pay in several installments for the work I do for you or for your purchase of my car, I can propose that we sign a legally enforceable contract. Now you are faced with the likelihood of real costs if you renege on the trust I have placed in you. Or if my fellow citizens have

entrusted me to enforce the law, they may have institutional devices available to remove me from office if I fail my trust. In these and many other cases, institutions help to secure your and my trustworthiness.

Learning to trust will depend on the success of trusting, which will turn on the trustworthiness of those we trust. Hence, we can imagine that enhancing trustworthiness in general will increase levels of trust, which will induce more productive cooperation. In this sense, trust can be enhanced by introducing devices, such as the law of contracts, to regulate relationships to make parties to them more trustworthy. More generally, the development of norms with sanctions and of other devices for social control tends to enhance cooperation and reduce the risks inherent in trusting others.\textsuperscript{12} Hence, the effect of institutional enforcement of trustworthiness may go well beyond making specific instances of trust reasonable.

Let us unpack this claim. How can trust be enhanced by enforceable contract (or by audits with the threat of sanction)? The contract or audit may protect a relationship against the worst of all risks it might entail, thereby enabling the parties to cooperate on less risky matters. This is a milder and more specific variant of the central argument in Hobbes's theory of political order. The threat of sanctions to protect each makes all better off. Hobbes's sovereign enforces order in the protection of life and property. Once these are secured, we have less reason to be defensive and we engage in productive investments and beneficial exchanges.

James Coleman describes an emergency spot loan from the Hambros Bank in London to a ship owner to let his ship leave port in Amsterdam. The loan was made instantly without security. Coleman says that it was made with nothing more substantial than the ship owner's intention to repay and the Hambros loan officer's belief in both the ship owner's honesty and his ability to repay.\textsuperscript{13} This claim cannot really be true. There was some threat of suit and court enforcement. Going through the courts to get enforcement, however, would likely have cost Hambros something even if the ship owner were legally forced to pay in full. Suit may often be a necessary background possibility to get both sides to come to reasonable agreement, because both sides will want to avoid what, to them, would be the deadweight losses of legal action. But legal action can keep the scale of losses from misplaced trust substantially lower than the amount of money involved in the initial trusting agreement. The possibility of occasionally having


\textsuperscript{13} Coleman, p. 92.
to go to court can therefore be counted as a relatively slight cost of doing profitable business.

Most parties to Hambros loans likely prove to be fully trustworthy to repay. But the background possibility of sanction greatly enhances this trustworthiness and justifies Hambros's general risk taking. We might say that the use of courts, when it happens, is the centrally important consideration to Hambros. But this is likely false. What is centrally important is that the existence of the court sanction changes relations drastically across the board, making Hambros's relations with virtually all potential borrowers more cooperative than they would otherwise be.

In deciding whether to trust someone, we can often avail ourselves of extant institutions. Hence, we can make more credible commitments on some matters than on others, depending on what institutions are available. Until recently, Italians could make lifelong commitments to marriage. The commitment was limited. For example, it did not entail marital fidelity or even cohabitation. It merely guaranteed that neither partner could remarry while the other partner still lived, and this implied that children with a substitute spouse were illegitimate. Today, Italians' commitments to marriage often rival Americans' in their lassitude. Among some groups, marriage is by whim—and only so long as the whim lasts. The threat of AIDS may now do more for marital fidelity than either law or love does in many communities.

Given our institutions, what can we consider trustworthy? Remarkably, economic institutions such as manufacturers and dealers are often trustworthy, at least in their dealings with their customers. (They might not be trustworthy to the larger society insofar as they might readily pollute or discriminate by neighborhood.) Of course, when a manufacturer or dealer faces bankruptcy, its usual reputational and legally induced incentives to be trustworthy can deteriorate badly. When future expectations collapse for any such reason—the previously trusted party suffers from radically changed circumstances—there is less reason to expect trustworthiness. 14

Similarly, governments may often be quite trustworthy in carrying out the laws, even though they might be very untrustworthy in designing the laws to be carried out. In both these cases, institutional control devices are often powerful and well directed. These devices have no match in professional contexts, although control of professional behavior may be changing toward sterner regulation. Similarly, economic relations between individuals are relatively well governed by

in institutional controls. Other relations are typically much less well governed. Some areas are coming under increasing institutional oversight, as are some gender relations and some intrafamilial relations. But still, contract law, for all its difficulties, seems almost like an ideal type in comparison to many other institutional controls over individual commitments.

Many social critics complain of the seemingly increasingly economic and material focus of modern life. The complaint is that we have let economic concerns and material welfare displace other values. Barry Schwartz argues that, in their misguided scientism, the disciplines of behavior theory, economics, and sociobiology may give an accurate picture of things as they are, of human nature in our circumstances. But, he supposes, this is because these disciplines have contributed to and helped justify the conditions that foster pursuit of self-interest.\textsuperscript{15} The complaint is surely misplaced. The academic disciplines of the social sciences may have made a big difference in the self-understandings and values of social scientists, but it seems implausible that they have so radically affected the values of others in our societies. The social sciences are better at discovering where things have gone than in pointing or leading the way.

There is a deeper problem that may be the greatest constraint on what our values can be. Our institutions enhance trustworthiness and, hence, the value of trust far better in economic than in noneconomic relations. One might suppose that this follows because these institutions are deliberately designed to work that way, whereas our social conventions merely arise and grow and are generally not meaningfully said to be intended. But institutions and organizations also grow in unintended ways, and they need not be true to any designers' intentions for them. More crucial to the difference in the economic and noneconomic constraints may simply be the fact that the former are easier to assess and easier to build into strong expectations—and therefore easier for institutions to enforce. There is great trustworthiness in contracts because performance is easy to assess and enforcement is relatively easy; there is far less trustworthiness in marriage in many societies and times, because performance is too hard to measure to make enforcement work. Not surprisingly, even in marriage economic interests can be monitored well enough to be governed by contractual agreement, and there may be a growing trend toward prenuptial contracts to cover financial matters.\textsuperscript{16}


\textsuperscript{16} “Prenuptial Pacts Rise, Prenuptial Trust Fails,” \textit{New York Times} (November 19, 1986). This is a less revolutionary change than one might think. In many times and places, national and familial interests have dominated some marriages, and powerful
Incidentally, it is arguable that, if contracts become as shaky as marriage, then our society will be in danger of collapse. That is to say, ordinary contractual regulation may be more important to social order in a complex society than is the enforcement of any particular possible convention on sexual and familial relations.

SOCIAL CONSTRAINTS

The social constraints of conventions and norms are often reinforced by more formal institutional constraints, including law. Such reinforcement is not typically necessary, but it is likely to be a natural move. It can happen from the common but specious supposition that what is ought to be—and therefore ought to be legally or otherwise enforced. Or it can happen because those who back a convention or norm can get it reinforced by law. Consider two striking conventions that motivated the fictional lives of Ukifune in Lady Murasaki’s *Tale of Genji* and Violetta in Giuseppe Verdi’s *La Traviata*. The first of these worked forcefully without need of law to back it. The second was made especially forceful by legal constraint.

First, consider the unusually powerful social convention that governed the commitments of the beautiful young lady Ukifune in eleventh-century Japan in the twilight of the world of the shining Prince Genji. Ukifune left her worldly existence by having her head shaved, taking vows, and entering a nunnery, after which she could never go back to her previous world. Cutting off hair that it had taken her a lifetime to grow to six feet long was, in that society, a step that was visually and symbolically irreversible. Ukifune therefore had little choice other than to live up to her sudden commitment to her religious vows into the distant future, although that meant a personally and culturally impoverished life compared to life at court. As distressed as everyone around her was that she had become a nun, no one thought she could renege and return to society. She might grow to hate her life as a nun and wish she had never made such a commitment, but, short of suicide, she was virtually stuck with it.

Ukifune chose to leave her world because she had two potential lovers—the two most desirable men in her world—and she preferred

---

Institutions have stood behind those interests. The principal change in contemporary conditions may be that far larger fractions of marriages put substantial economic interests at stake. Schwartz argues that, once satisfaction of personal interests “becomes the goal of friendship and marriage, it destroys these activities as distinct, organized, and coherent” (Schwartz, p. 269). Unless he means the only goal, or these terms are defined in odd ways, this is utterly implausible. Marriages must often have strong, even predominant elements of satisfaction of personal (and other) interests. Are they then destroyed as a “distinct, organized, and coherent” activity?

to have none rather than to have to choose between them. Yet this woefully weak person made a stalwart commitment. At first, she intended to leave the world by dying, but she failed and was rescued by chance to a nearby nunnery. Becoming a nun was then purely opportunistic. The only actual commitment she had was to leave her prior world. If she could not be dead, she could at least be a nun even though she might have little commitment to any of the ostensible beliefs of Buddhism.

Under the social conventions of her time, Ukifune's was a virtually irrevocable commitment. In our society of radically looser conventions, you and I cannot so readily constrain ourselves as Ukifune did—our shaved heads might be nothing more than a frivolous style of the moment—and we cannot be fully believed if we assert our undying commitment to a particular religious creed or to any other purpose. We can be believed to say we are committed in this very moment—but we cannot be fully trusted to stay committed into the distant future. The most assertive believers among us are all too often the most fickle. We are like Alberich without magic to control us. We might say, “Love, I curse you forever,” but we could not be sure to live the sentiment. Ukifune could and did—because she had a whole society to back her commitment.

In the second instance of social constraint, convention may be powerful, but it is typically heavily overlaid with law in modern societies in governing what for many people in many times and places must be the single most important problem for trust: marital fidelity. Convention matters because large-number stability in the behavior of relevant others or other types can be a reason for expecting more of the same behavior. If virtually all spouses are faithful, I can be more confident of my own spouse. Large-number regularities can play a stronger role: they can affect one’s incentive to act as others do. My expectations of your behavior may turn in large part on my expectations of behavior of people like yourself in our society.

I might suppose that the best norm for marriage is to be utterly faithful until death do us part. If all followed that norm, I might firmly believe, ours would be a better world. Suppose the woman I wish to marry believes the same. We therefore publicly vow to be faithful forever. Should we believe each other? Perhaps. But perhaps only in the sense that we really are committed at this moment. Whether we will still be committed in a few years may depend very much on the laws and norms of our society. “If the belief that marriage is just a devious route to commercial sex becomes widespread, then,” Barry Schwartz argues, “individuals will have no choice but to view it that way themselves.”

have to live with the larger society's coordination on a particular pattern of expectations. For example, I cannot reasonably expect you to be dramatically different in your long-run commitments than others are.

In Verdi's *La Traviata*, Germont wishes to discourage the former courtesan Violetta from continuing her life with his son, Alfredo. Violetta says that, alas, she can love no one but Alfredo. Changing his tack, Germont points out that "men are often fickle." Violetta involuntarily says, "Oh God," and Germont, given the opening, demolition her hope. He sings:

One day, when time has dispelled the charms of love,
tedium will set in quickly.
What happens then? Think—
The deepest feelings can bring you no balm,
since heaven has not blessed this union.19

The blessing of heaven was, of course, to be secured from earthly institutions that would then protect Violetta's claims on Alfredo by, at the very least, making it illegal for him to marry anyone else. Violetta understood that restricting one's freedom might be necessary to securing desirable ends. Of course, her greater concern was to restrict someone else's freedom to secure desirable ends.

Leo Tolstoy reports the code of honor of the aristocracy of nineteenth-century Russia.20 Given the prevalence of this code among his class, Count Vronsky in *Anna Karenina* was virtually incapable of convincing anyone he would behave otherwise toward anyone's beautiful wife than he behaved toward Karenin's Anna. In this respect, trust, like expectations and knowledge, is in part inherently a social construction. Even if they were trustworthy, Vronsky and Alfredo would have no way to convince others. Again, the problems of commonsense epistemology overwhelm people in Vronsky's and Alfredo's societies—what one can communicate at all may be socially constrained or even determined.

If divorce is impossible and infidelity is severely sanctioned, we who wish to swear fidelity until death may find it quite easy to trust one another. If infidelity and opportunities for it are rampant, however, we may be unable to trust one another with anything short of an act of faith. In many contexts your capacity to make a commitment about

how you will behave depends on the normal expectations of behavior. You cannot easily establish a commitment to a much more demanding standard of behavior. Genuinely normative commitments may therefore depend on or interact with incentive structures.

If love were not fickle and marriage were generally faithful, half or more of the world's great literature would not exist. But we know better. Lady Murasaki began one of her liaisons (ostensibly with Fujiwara no Nobutaka) with a dismissive poem. In her own Memoirs, she prefaced it as follows:

Someone who I heard was in love with the daughter of the Governor of Omi said:

“I have no other love but you!” He went on insisting and I became annoyed:
Plover crying to his mate
On the waters of the lake,
I would as lief
You did not cease
To call at many ports.21

In an extremely small, tight-knit society, every peccadillo might instantaneously become public knowledge, and there might be very little infidelity—or infidelity might be relatively tolerated, even more than it was in medieval Japanese society or in the France of the last three Louis. In contemporary, loosely connected societies, monitoring must be much less reliable and infidelity much less often discovered. But for this very reason it may be far more suspected, and the discounting of one's partner's future affections must often lead to the readier decay of one's own.

Why do informal social conventions work to secure our capacity to commit to trustworthiness? In part for the reasons that organizational devices work. Perhaps the best way to account for such workability is to characterize social conventions that do work. Very often, the rules that can be made to work well are especially economical. They can be simply stated, breaches of them and adherence to them can be easily monitored, and the cost of invoking them must bear heavily on the rule breaker and lightly on any sanctioner. In addition, they must be relatively stable, so that their effects can be predictably expected into the relevant future.

For example, Ukifune wanted to commit to leaving her world, the world of the court in Heian Japan. She did so by taking vows, because this excluded her return to the secular world. First, the rules

of her commitment are easily stated: she must live in a convent and not range into the larger world. Second, her actions were extremely easily monitored. If she left, her absence would be instantly noticed by the small group of her fellow nuns. And her presence elsewhere would be instantly seen as wrong even by those who knew nothing of her, because they would be able to see that her hair had been shorn. Third, the sanction of any attempt on her part to return to the court would be essentially of no cost to most of the courtesans, courtiers, and other hangers-on at the court—they would almost all have coordinated on shunning her. But that sanction would have been devastating to Ukifune if she had really been intent on returning to the court. Finally, the rules governing her behavior after taking vows could firmly be expected to last long beyond her life. They were stable.

Ukifune almost perfectly arranged her commitment to leave the court. One might suppose she made a very bad choice, but she would evidently have disagreed at the time of the choice. Or one might even suppose that she could have left the court in a less degrading and stultifying way. But it is not easy to imagine what more attractive devices she had available in her close-knit society.

For many other forms of commitment through fitting one's desired future actions to social constraints, the use of social pressure may loom large, as in Ukifune's case. Even while her aristocratic identity was not yet known to the nuns and priests, they attempted to pressure her not to take so drastic a step. Her greatest struggle in the short term was to overcome that pressure and to get the priest to shear her. She, more than most of us who might analyze her problems, knew very well the weight and implications of social pressure, and she set herself up to use such pressure against her future self, to guarantee her commitment. Such social pressure may generally work best when the relevant rule works at the small-group level, that is to say, in a context of thick relationships. Ukifune's constraints transcended her small community at court, because everyone in Heian Japan would have recognized that she had left the world.

If Ukifune had merely done something shameful, something that was not instantly visible to one and all, she could still have been excluded by the small world of the court, where her transgression and her identity would have been known to virtually all. Her problem was evidently that she did not merely wish to absent herself from the court but that she wished to block herself, in a weaker moment, from returning later, and she wished to block others from attempting to draw her back. She found a device that made herself trustworthy into the distant future and that simultaneously made others trustworthy not to try to change her mind.
SELF-TRUST

An instructive way to characterize the problem of commitment is as a problem of trusting oneself, of self-trust. Can one trust oneself? As with ordinary trust of another, this is not a singular question. It involves at least three parts. What can I trust myself to do? What can I trust myself to do? And what can I trust myself to do? I may be morally certain that I will live up to some commitments and much less confident that I will live up to others. I may know very well that, under the influence of camaraderie and drink at a party, I will suffer weakness of will and will stay too late and therefore fail to get something done that I have committed for tomorrow. You may know yourself very well and therefore know that you will not do the sensible thing at certain moments. That is, you would not trust yourself in that respect.

You might not know whether to trust yourself in some contexts. For example, people sometimes unexpectedly discover themselves to be courageous or cowardly only when faced with a physical threat to themselves or others. And politicians who have never faced serious bribery and who take a strong position against political graft might be surprised to find themselves willing to accept bribes when offered. We cannot be sure we know our limits when we have never tested them.

Incidentally, even the rather formidable and incomparably insightful Lady Murasaki, who seemed to comprehend everyone, plausibly could not fully trust herself. In the poem quoted above, she told Fujiwara no Nobutaka that she would sooner he continued to call at other ports—implicitly, to leave her alone. Yet she soon was heavily involved with him in what is thought to have been her one real love affair. Although he had other wives, she even married him. Her seemingly tough character did not withstand Nobutaka's energetic blandishments.

Suppose I know my limits in some respect and now wish to make myself fulfill some commitment. I can do various kinds of things. I can avoid the party that would wreck my evening's chance of getting some task done as promised. Or I can go with someone who will reliably drag me away before the third glass of wine or the endless conversation on current politics. I can use the devices canvassed by Jon Elster and Thomas Schelling. For some problems, I can take advantage of formalized institutional arrangements to sanction me

22. See also Partha Dasgupta, "Trust As a Commodity," in Gambetta, pp. 49–72, p. 54.
into fulfilling my commitment. If we stop to think through why we fulfill many commitments, we are apt to find such devices at work. We may sometimes fulfill them because we are dealing with someone whom we especially value, but often we do so because there would be bad consequences for ourself, for our reputations or our future prospects, if we do not fulfill them. Even when our commitment is literally to ourself alone, and not to others, we may often find that we enforce it by setting ourself up for social sanction if we fail. We talk about our thrice-weekly workouts, perhaps far more than would interest anyone, in part because we then face embarrassment if we fail to work out.

The self-manipulation to get ourself to the gym is a trivial case. Consider a much less trivial case, a woman who is Catholic but who sees around her a majority population who are not Catholic and a majority of supposed Catholics who are less than faithful. Suppose she fears that she herself might slide away in the future—after all, she has seen others do so, including her husband—and she fears that her children may be seduced into the secular world. What can she do to secure or strengthen her future commitment to the faith? She could join a religious order, but that might be too drastic, and anyway she does not wish to abandon her children to the secular world.

Short of such a drastic move, the woman can do something that gets her more involved in the church and that constructs about her a society in which she would lose face if she relaxed her commitments. For example, she could take up the cause of abortion—not because she was hostile to abortion but because, by identifying herself with the church's antiabortion movement, she would identify herself more closely with the church. She is like Ukifune. She relocates herself in the world of Catholic political activism not because she shares the activist agenda but rather because being in that world keeps her safer from the secular world of declining American Catholicism. She fears for her continued faith and plumps for rigid action on abortion to lock herself in. She throws herself in the way of fate, which, as Lady Murasaki says, can change her desires.\(^{25}\) In a more securely Catholic society, the fear of falling away from the faith might be less compelling because there are natural reinforcements, and there might be less urgent need to become deeply involved in the political activities of the church. But in her society, she may need to construct her own reinforcements to keep her faith.

CONCLUDING REMARKS

All of the devices for commitment discussed above, other than Alberich's act of pure willing, involve contriving to make it the individual's

\(^{25}\) Bowring, p. 234.
interest to be trustworthy. Or, plausibly more often, they involve choosing to commit to doing what is the individual's interest to be trustworthy to do.

Interest is one of the best and most useful of internal motivations. It leads us to and through much of the best that life has to offer. To suppress interest, on occasions when it leads us in ways we might consider wrong, is to put ourselves at war with ourselves. In waging that war we are supported by social controls based in our own long-term interest. A strong network of laws and conventions is needed to make any kind of behavior reliable if it is likely to conflict with powerful considerations of interest. Neither the law nor conventions were available to secure faithfulness from Alfredo, although love might have sufficed for him. All too often we have neither law nor convention nor love to make us trustworthy, and it is often only the interest we have in maintaining particular relationships that makes it in our interest to be trustworthy.

Trustworthiness may be inherently moral in part for at least some people. That many accounts of trust are really accounts of trustworthiness therefore suggests that the moralizing of trust might be more reasonably seen as a moralizing of trustworthiness. Certain standard moral theories, such as that of Kant or various virtue theories, could readily elevate trustworthiness to moral status. Surely there will be fewer theories, if any, that moralize trust itself, although writers whose focus is ostensibly on the phenomenon or idea of trust often do spuriously moralize it.