

Victoria McGeer

Developing Trust on the Internet

Abstract: Does the Internet provide an environment in which rational individuals can initiate and maintain relationships of interpersonal trust? This paper argues that it does. It begins by examining distinctive challenges facing would-be trusters on the net, concluding that, however distinctive, such challenges are not unique to the Internet, so cannot be cited as grounds for disparaging the rationality of Internet trust. Nevertheless, these challenges point up the importance of developing mature capacities for trust, since immature trusters are particularly vulnerable to the liabilities of Internet trust. This suggests that Internet trust can only be rational for those who have developed mature capacities for trust. But that suggestion ignores how trust on the Internet may also facilitate the development of such capacities.

0. Introduction

My aim in this paper is to consider whether the Internet provides an environment in which it is possible to engage in rational trust. However, before proceeding with this inquiry, a number of caveats are in order. For the topic is overly broad in at least three different respects: first, with respect to the *types of Internet exchanges* that can be examined with this question in mind; secondly, with respect to the *types of user groups* that provide relevant contexts in which issues of trust can be raised; and finally, with respect to the kinds of *actions and attitudes* that theorists could reasonably have in mind when they speak of Internet ‘trust’. My own discussion will be constrained in the following ways. I will focus on the kinds of exchanges that are interpersonal in nature—hence, that occur within, and so presuppose, a sense of Internet community built upon repetitive contact amongst a group of familiar users. I will also focus on a particular kind of interpersonal trust, ‘friendship trust’, where the primary mode of vulnerability is psychological or emotional. (Hereafter, all references to ‘trust’ should be understood as references to trust in this sense unless otherwise stipulated.) Finally, I will come to interpret my question about the rationality of Internet trust as composed of two parts: First, and most obviously, it involves inquiring into whether trust in this medium can be reasonably based given the kind of contact with others the Internet allows; but, secondly I claim, it should also involve an investigation into how interpersonal exchanges in this medium could enable the development of our capacities for rational trust, whether on- or off-line.

This paper will proceed in three sections. In the first section, I consider reasons that argue in favour of supposing the Internet is a particularly bad medium for ‘rational’ trust, by which I will mean initially, a particularly bad medium for

reasonably extending our trust to others. Of primary concern here is the sort of evidence available to us of the trustworthiness of others so encountered. For on the Internet, it's not just that individuals may hide importantly relevant features of themselves; they may actively mislead us in ways that invite our confidence. Hence, the decisions we make about when and how much to trust others are likely to be ill-informed or, worse, actively manipulated by them. To sharpen this discussion, I will give close attention to a particular case of Internet trust and betrayal. My aim here is to motivate conceptual distinctions that set the stage for arguing, in section two, that we need a more nuanced approach to the problem of Internet trust than such general concerns about the medium allow. In particular, I argue that the difficulties of Internet trust are importantly dependent on—i.e., magnified or decreased by—the maturity one brings to trusting relationships in general. The more mature the truster, the more capable they will be of reasonable trust, even on the Internet—i.e. trust that takes account of the limitations inherent in the situation. Hence, there is no good reason to argue against the rationality of Internet trust in general. However, this does not yet say much in favour of relationships of Internet trust. In the third section, I redress this imbalance by touching briefly on relationships of trust among virtual persons in the context of virtual communities. My aim here is to stress the developmental potential for us as trusters through engaging in relationships of Internet trust.

1. Julie's Tale: Does the Internet Facilitate Irrational Trust?

The following is a true story of Internet trust and betrayal.¹ Julie was a deeply disabled older woman living in New York City who could push the keys of a computer with her head-stick. Highly limited in what she could do in the physical world, Julie found her *métier* on the Internet. Off-line, she was cabined, cribbed and confined. On-line, she was able to lead a rich social life in keeping with her expansive personality. Warm-hearted, talkative and caring, Julie soon became a popular member of a New York Internet conference or chat-room composed of like-minded women with whom she could fraternize on equal terms. Although she did not hide her disability from others, on the Internet it presented no physical or emotional barrier to be overcome. If anything, the handicaps she bore with patience and good will became an inspiration to others, making her all the more likeable. In addition, Julie was an extraordinarily good listener to other women's difficulties. She was perceptive, articulate, thoughtful, full of good advice, and seemingly endlessly patient as many women whom she met through the net poured out their troubles and concerns. She became a solid friend to many, who felt that their lives had changed for the better through knowing Julie and from taking her advice.

But this story has an unhappy ending. After several years of participating

¹ Julie's story dates from 1985 and is documented in Stone 1991. Thanks to Helen Keane for giving me this reference.

in this Internet community, one of Julie's devoted admirers decided she wanted to further their friendship in person. Tracking Julie down in her New York apartment, the woman discovered that 'Julie' was no Julie at all. In fact, she was a man—a middle-aged psychiatrist who, besides being male, wasn't disabled at all. The woman was outraged, and after she made this fact known to the wider conference, the conference itself was deeply shaken. Reactions varied from a kind of resigned amusement all the way through to a feeling of total betrayal, especially amongst those who had shared their innermost feelings with Julie. One woman reputedly said, "I felt raped. I felt that my deepest secrets had been violated". Worse, those who had made genuine gains in their personal and emotional lives felt that these gains were predicated on "deceit and trickery", hence were stripped of any value and should be repudiated.

Although we haven't yet heard Julie's side of the story, this tale makes vivid why friendship trust on the Internet is a risky business. Such trust, as I said before, involves relying on someone in ways that make one *psychologically or emotionally* vulnerable to them rather than simply materially vulnerable, although the psychological and the material are often intertwined in more or less complicated ways. Still, my focus is on the kind of psychological vulnerability that comes from trusting others with *ourselves* rather than with our credit card numbers, or banking information, or to discharge contractual obligations, and so on. This, after all, is the way in which Julie's friends and admirers felt betrayed—in the way they understood themselves to relate most intimately and personally to another human being.

But now why did Julie's friends and admirers feel betrayed? The obvious answer is that Julie was not who she pretended to be. 'She' was a made-up character, nothing but a fictional being, a mere puppet masking the true identity of someone altogether different. And the Internet provides ample opportunity for such deceit. For instance, people can hide their real selves in whole or in part, masking facts about their gender, ethnicity, age, appearance, health etc.; they can assume a variety of different identities with ease, becoming more than one phoney self; or they can even band together in real life in order to enact a single phoney self. This is consequent on two features of Internet contact that give individuals inordinate amounts of control in what information they convey about themselves to potentially trusting others: first, it is primarily a *text-based medium* of interaction (though video and voice contact are now also used with more frequency); and secondly, it is primarily a *uni-dimensional mode* of interaction—i.e., it is normally only text, a single medium for conveying the sort of information on which those who trust must rely for making judgements about the trustworthiness of others.

Consider now how both of these features interact to play into a potential trickster's hands. For instance, the fact that Internet contact is both textual and uni-dimensional means that we cannot depend on our usual ways for cross-checking the information we receive about others based on what they tell us. Ordinarily, this involves our own observations of them based on their real bodily presence: their looks and manner, not only in relation to us, but in relation to others with whom we can see them interact. Importantly, such observations

have not only a cognitive component, but also an immediate visceral/ emotional component. We are often sensitive to subtleties of tone and manner that can't be expressed in words: We find ourselves reacting positively or negatively to others based on subliminally detected bodily cues. And we often are wise to trust our instincts under these conditions, since they constitute a reasonably accurate early warning system. For instance, researchers have established that subjects can become viscerally or bodily aware of a bad or dangerous situation well before they can articulate their concerns in a cognitively explicit way (Bechara et al. 1996; 1997).² Internet communication thus deprives us of the bodily information we need for this early warning system to operate effectively.

Furthermore, it seems that text-based interactions are particularly seductive in character when there is no sensory check on how to interpret what others say. For even when we are not actively misled by their words, our imaginations tend to roam freely over textually underspecified details. This imaginative freedom may be further encouraged by an unwitting hubris— a conviction that we have more judgemental control over the emotional and cognitive effects of 'mere words' than experience suggests. Consider, for instance, the powerful illusion created by the computer program ELIZA. ELIZA was developed by Joseph Weizenbaum at the MIT Artificial Intelligence Laboratory in the early 1960s. In response to text-based input from ordinary human users, ELIZA would analyse the syntactic structure of the text, and respond with a seemingly appropriate question or comment, modelled on the kinds of questions and comments made by psychotherapists. Instead of being bored or tipped off by the fairly limited style of response, people found the illusion of agency—indeed, the illusion of caring, concerned agency – remarkably seductive. Weizenbaum reports, for instance, that when he first introduced people to the program, they would spend hours with it discussing their personal problems, unwilling to believe that there was not a real agent behind the seemingly concerned and attentive ELIZA (Weizenbaum 1976). Even now, there are Internet websites where individuals can go to pour out their troubles to ELIZA, seemingly unfazed by clear statements that ELIZA is a mindless program. They either don't believe it, or they don't care—the illusion of concerned agency is more than enough for a good conversation, at least of a particular therapeutic sort.

The more general conclusion is this: text-based interactions are a powerful stimulant to our imaginations; they have a persuasive power all their own, and through them, we are remarkably susceptible to projecting on to the originators of that text (be they human or robotic) whatever qualities we detect in the text. Our judgemental distance is often much less than we imagine it to be, particularly when these interactions involve emotional or sensory contents. This makes us particularly prone to illusion in the domain of textual communication, especially when this is the only medium available to us.

To summarize the problem with trust on the Internet thus far: It seems that the Internet is a particularly bad medium for fomenting good, stable, non-delusory relationships of trust—hence, what we might call 'rational' or 'reason-

² For further philosophical discussion of how our emotional reactions can appropriately shape our reactions to the world, see Jones 2003.

able' trust—for two reasons. The first I will call the *Proteus factor*, after the Greek sea-god fabled to assume various shapes. The Proteus factor refers to the ease with which people can hide who or what they really are on the Internet. This may be done explicitly as in Internet role-playing games (I will come back to these later in my paper), or it may be done implicitly, with individuals actively working to deceive others. The second reason the Internet is a bad medium for fomenting rational trust involves what I will call the *Eliza factor*: This refers to the ease with which people can be seduced by text – particularly interactive conversational text—viewing it willy-nilly as a kind of window on the soul of the agent who produces it, rather than (more cautiously) as a kind of mirror that partly reflects their own imaginative projections. In other words, people often see what they want to see, especially when their 'seeing' is solely mediated by interpreting a disembodied string of apparently trust-conducive utterances.

This is a strong conclusion, and while I will not ultimately endorse it, I also don't want to minimize the strength of the considerations supporting it. These are liabilities for trust that are particularly pronounced on the Internet. However, it is also important to note that computer mediated communication is not the only way in which these two factors can play a role in promoting illusions and so undermining the conditions for rational trust.

Consider the unhappy story of *Cyrano de Bergerac* (Rostand 1959). A large-nosed romantic, Cyrano considers himself too ugly to win his cousin Roxanne's love. Consequently, he spends years languishing by her side in the role of trusted friend and confidante, never confessing his deep love for her. Roxanne, meanwhile, is on the lookout for true love, and after a few fleeting encounters with Cyrano's dashing and handsome friend Christian, she thinks she has found the perfect mate. However, all is not well for these would-be lovers. It turns out that Christian does not have the kind of poetic soul that Roxanne requires in a lover. He can't win her by good looks alone; and Cyrano, whose soul is sufficiently poetic, assumes she will reject him because of his unfortunate appearance. So the two men team up: Christian becomes Cyrano's physical proxy, a mouth-piece for Cyrano's words that eventually win Roxanne's undying devotion. But the question is, with whom is she in love? Not with Christian, whose personality is kept hidden from her, and not with Cyrano either, whose position as loyal and avuncular cousin is a mainstay in her life. The object of her love is, in fact, a fiction—an amalgamation of the two men. It can't survive in the real world, because even though Roxanne physically knows and interacts with both Christian and Cyrano, the man who she really loves is a fantasized projection of her own desire, supported by Cyrano's Protean deceit and her own Eliza-like capacity to be seduced by what becomes largely text-based communication (the many letters that Cyrano writes to her under Christian's name).³ The story ends in tragedy, but it is a kind of pathetic tragedy in which one really feels that a little more honesty at certain crucial moments would have at least given Roxanne some capacity to make a sensible decision about whom to love and trust, given the realities of her situation.

³ For a compelling philosophical discussion of the impossibility of Roxanne's love, see Campbell 1997.

Of course, this story is only fiction, but it serves to show that the Internet only facilitates the kinds of practices of deceit and self-deceit that can occur in the physical presence of others, where presumably one can rely on other modes of interaction to check and qualify the judgements one makes about them. More sobering examples include all those real-life confidence tricksters who make a practice of relying on their physical attributes—honest appearance, convincing tone of voice, smooth manner—to dupe their victims, however normally savvy such victims might be. Hence, multimedia trust—trust that is developed through various media of interaction—may help reduce the likelihood of that trust being grounded on illusion, but it can't eliminate this possibility altogether. This is one cautious thing to be said against the detractors of Internet trust. But I think something more powerful can also be said in its defence. In order to do this, I return to the story of Julie.

2. Julie's Tale Revisited: Immature Versus Mature Relationships of Trust

In some ways, it is easy to understand and sympathize with the sense of betrayal that Julie's friends and admirers in this Internet community felt when her true identity was revealed to them. Julie had deliberately lied to them about who and what she was. Still, I hope there is also some discomfort with the reactions reported by some of these women. Recall that one reputedly claimed to have felt 'raped', and another, that the positive gains in her life were worth nothing since they had been based on 'deceit and trickery'. Of course, these extreme reactions were not universally shared. Other women, also friends of Julie's, confessed to feeling little more than astonishment followed by humorous resignation as if something like this might well have been expected—not condoned, mind you, not invited, but reasonably anticipated.

What accounts for this range of reactions? The straightforward answer is that these women differed in the *amount* of trust they had invested in Julie and in their relationships with her: i.e., those who were not that trusting to begin with felt relatively less betrayed than those who were more trusting. While I think there is some truth in this, a mere quantitative analysis is not nuanced enough to do justice to the real variety of possible explanations. For instance, it doesn't distinguish between two sorts of cases: (1) women who were less trusting in this situation because they were generally less trusting, i.e. generally less capable of trust in others; and (2) women who were 'less' trusting in this situation because they were generally more capable of trusting well, hence of trusting appropriately (i.e., with appropriate sorts of expectations) relative to the situation. In other words, the wide range of reactions to Julie's revelation might partially be accounted for in terms of differences in the *quality* of trust extended by these women.

My point here is that betrayal is a complicated phenomenon. Obviously, when someone betrays another's trust, they have inflicted a harm on the one betrayed. But how that harm is experienced by the person betrayed and what

its consequences are will partially depend on that person's expectations and capacities with respect to initiating and maintaining relationships of trust. Thus, in contemplating Julie's story, one question we should want to pose is the following: what does it mean to trust well, to trust responsibly, to trust reasonably in someone over the Internet? And might the women who felt most undermined by Julie's betrayal be partially faulted in terms of the quality of their trust, rather than simply in terms of the fact of their trust? By this shift of emphasis, I do not mean to minimize Julie's own part in this story of betrayal; I simply mean to probe more carefully into the dynamics of trusting relationships.

I will begin this examination with an interesting remark Hubert Dreyfus makes in his recent book, *On the Internet* (Dreyfus 2001). According to Dreyfus, there is a sort of trust that is not possible on the Internet because it requires bodily presence—that is, it requires our having the capacity, as he puts it, to look one another in the eye and shake one another's hands. In Dreyfus's words, "The kind of trust that requires such body contact is our trust that someone will act sympathetically to our interests even when so doing might go against his or her own" (Dreyfus 2001, 70). In my view, this falls under the rubric of 'friendship trust', precisely the sort of trust that Julie's friends had invested in her. Consequently, Dreyfus's remark is clearly an overstatement: such trust is certainly *possible* on the Internet. But perhaps he means to suggest, in keeping with the worry voiced in Section 1, that friendship trust is *irrational* on the Internet because trusting reasonably in this sense requires that we be bodily present to one another. Otherwise, Dreyfus implies, we open ourselves to making a kind of emotional error. For, in his view, well-grounded trust, "must draw on the sense of security and well-being each of us presumably experienced as babies in our caretakers' arms" (Dreyfus 2001, 71). If we can have such a sense of security absent the embodied presence of another, then we must simply be fantasizing such a presence—reading into the cool voice of text the warm embrace of another's arms. Less poetically, we mistakenly suppose that the traits of others that are manifested through our textual interactions with them are grounded in their caring embodied presence. Hence, we delusionally expose ourselves to the possibility of deep and damaging betrayal.

I think Dreyfus gives a good diagnosis of why some of Julie's friends might have been so badly undermined by the discover of her real identity. They trusted in her to provide just this kind of security in their lives. She was the maternal figure to whom they could turn for the absolute safety of a genuine, albeit imaginary embrace. Her physical embodiment in the real world thus really mattered to them—mattered in the sense of sustaining their fantasy of who lay behind the surface exchange of text. I will call this kind of trust 'security trust', and I agree with Dreyfus that such trust can never be well-grounded on the Internet, largely because of the Proteus factor (the malleability of Internet identity).

However, I think there is a further question to ask about security trust—viz., is it the sort of trust that can be faulted more generally for its irrationality? In other words, is it the sort of trust that rational, self-standing adults should be seeking whether on the Internet or off it? My worry is that in so far as we trust this way as adults, we're looking for the wrong sort of thing in our trusting

relations with others—viz., a kind of security that relationships among self-standing adults cannot reasonably deliver. Consequently, relationships governed by these expectations are particularly vulnerable to breakdown; and when they break down, generate experiences of betrayal that are particularly traumatic. The problems of security trust may thus be exacerbated on the Internet, but they are not unique to the Internet. They stem, more fundamentally, from the fact that security trust is an immature form of trust to which we may all be attracted, but which ought to be guarded against, not by abandoning genuine relationships of trust—that *would* be a kind of madness—but by becoming more mature both in our understanding and in our enactment of such relationships.

To flesh out this claim, I want to spend some time considering the differences between mature and immature relationships of trust, and in particular, the kind of dependency or vulnerability that characterizes the quality of trust in each. Following Dreyfus's lead, I will begin by characterizing what many refer to as 'infant trust'—the sense of security we first experience in our parents' arms.⁴

Infant trust, we may say, is distinctive because of an infant's utter dependency on, and hence vulnerability to, others. As Lars Hertzberg says, "the human infant is not ... an independently intelligible living unit, and not simply because of the physical cares which he must receive from others, but because the sense of his activity depends on the way in which it is interwoven with the activity of other" (Hertzberg 1989, 316). The developmental psychologist, Jerome Bruner, has called the sense-making structuring of the infant's activity, 'parental scaffolding'. The idea behind it is that babies come into the world without much capacity for self-maintenance, still less with the capacity for self directed thought and action and, hence, for self-determination. Nevertheless, they have impressive capacities for imitation and, in particular, selective imitation, first, of the facial movements of their caretakers, then of body movements, and finally of actions with objects in their shared environment. These mutual imitation games, delighted in by babies and parents alike, are the primary means by which infants identify themselves as like others and so, eventually, as persons whose thoughts and actions belong to the kind that persons produce. They are also the primary means by which parents mould children to react, think, and feel about things as persons do. As the psychologists Meltzoff and Gopnik remark,

"... mutual imitation games are a unique and important constituent of early interpersonal growth. Adults are both selective and interpretive in the behaviour they reflect back to the child. They provide interpretive imitations to their infants, reflections that capture aspects of the infants activity, but then go beyond it to read in intentions and goals to that behaviour. ... This, in turn, leads the infant beyond his or her initial starting point. Likewise, selected actions, especially those that are potentially meaningful in the culture, will be reflected back [to the infant] more often than others." (Meltzoff/Gopnik 1993, 349)

The dependency the child experiences in the hands of the adult is thus the

⁴ The following discussion draws substantially on the ideas I explore in McGeer 2002.

dependency of ‘self’ constitution. The parent literally makes it possible for the child to define and understand itself in social space, which is a space at the same time created by the parent. The child’s capacity for self-determination is thus, at this stage, taken on by the parent— eventually, of course, in order that the child can develop an independent capacity for self-determination. Paradoxically, then, self-determination must begin with other determination: the child becomes an agent by having its agency enacted by another. Now, is the bond in this relationship a bond of genuine trust?

There are good reasons to call it trust, but reasons not to as well. Trust, as many have pointed out, is not mere reliance, but reliance that importantly involves the mutual recognition of personhood. As Richard Holton claims, “Trusting someone does not involve relying on them and having some belief about them: a belief, perhaps that they are trustworthy. What it involves is relying on them ... and investing that reliance with a certain attitude.” This attitude we normally take only towards people. As Holton elaborates, “when the car breaks down we might be angry; but when a friend lets us down we feel betrayed” (Holton 1994, 67). Holton never fully clarifies why we adopt such different attitudes towards the things we rely on as opposed to the people we trust; but it seems clear that it must have to do with our expectations that others’ behaviour towards us will be governed by their seeing us as persons and, specifically, as persons who treat them as individuals capable of acting in a trustworthy way.⁵ Objects don’t do that. Trust thus involves, minimally, a complicated Gricean structure of person-specific recognition and acknowledgement. It involves (a) our acknowledging others as sources of self-determined action in their own right, with interests and desires worthy of respect; (b) others acknowledging us as sources of self-determined action in our own right, with interests and desires worthy of respect, and finally (c) each of us acknowledging that these attitudes are shared between us, and govern our actions and reactive attitudes towards one another. Without these attitudes, and their mutual recognition, we would be incapable of moral interactions (Strawson 1974).

Now, in the case of the relationship between parent and infant, it is clear that the infant’s development as an independent agent depends on the parent’s acknowledgement of the infant’s personhood. But the infant is not yet a person, in the sense that he is likewise capable of recognizing either the parent or himself as a person each in his or her own right. At best, the infant is capable of what Meltzoff and Gopnik describe as a kind of functional recognition: here is something ‘like me’, i.e., something that can be imitated and imitates me in return (Meltzoff/Gopnik 1993). So the infant is not yet in a position to trust the parent. But it is trust-in-the-making, and made only because the parent behaves *as if* the child trusts the parent—i.e. the parent acknowledges and acts towards the child as a person whose attitudes and actions towards the parent are not only self-determined, but also conditioned by the child’s recognition of the parent’s own personhood. In this way, the child develops the Gricean awareness

⁵ This theme has been sounded by a number of theorists working on trust. See, for example, Baier 1986; Jones 1996; Pettit 1995; Walker forthcoming.

of others and so comes to be the kind of being that can trust another, i.e. a being that is capable of full-blown adult trust (Hertzberg 1989).

With this kind of contrast case in the background, we're now in a position to ask: what is so distinctive about mature relationships of adult trust? To begin with, they must involve dependencies that are very different in kind from the dependencies of 'infant trust'. Since the trusting adult does rely upon the other, she is vulnerable, like the infant, to actions and attitudes outside of her control. Nevertheless, the adult relationship between the truster and the trustee is importantly symmetrical. Unlike the infant, the adult truster does not depend on the other either for self-determination or for maintaining the relationship between them as a relationship between persons. She is an autonomous person in her own right. The trust she gives is, therefore, genuinely chosen in that it issues from her own capacity for recognizing the relationship between her and the trustee as a relationship conditioned by mutual acknowledgement. And since such acknowledgement does not depend on the trustee's adopting the truster's role by treating her *as if* she were a person with self-directed thoughts and intentions, it follows that the thoughts and intentions the truster actually has must be acknowledged by the trustee if the trustee is genuinely to treat the truster as an autonomous person in her own right. (No parental make-believe is involved in such a relationship.)

The adult who trusts thus requires from the trustee something much more than the infant requires from the parent, and also something much less: The truster requires that her vulnerability to the trustee be recognized as the vulnerability of one self-determined person to another. It is thus a vulnerability based on interests, needs, and desires which are importantly the truster's own and to which she trusts the other can and will be sensitive, guiding his actions accordingly. But, of course, since the truster requires this kind of full acknowledgement from the trustee, she must give it as well—and thus be prepared for difficulty and disappointment. After all, the trustee is a person in his own right as well – with needs, desires, interests that are importantly his own. So, even with all the good will in the world the trustee may not be able to live up to the truster's hopes and expectations, either because he has misunderstood the truster's needs and desires or because his own needs and desires cannot be easily reconciled with the truster's own and cannot be given up without serious compromise. The difficulties of trust between self-standing adults are thus ones that stem from the need to recognize and negotiate autonomous interests. In particular, since in the adult case, the trustee does not take on the role of determining the truster's needs and desires, the trustee can betray the trusting adult in a way that the parent cannot betray the so-called trusting child. This is not to minimize what can happen to the child. On the contrary. The child can be profoundly and invasively damaged by the parent of whom it must rely. But because the parent takes on the role of determining a child's understanding of its own needs and desires, the child cannot experience, at least initially, the gap between self and other as a gap of potentially conflicting interests.

The bottom line, then, is this: The experience of conflicting interests and the threat of betrayal is much more salient in relationships of adult trust. As

we develop, we become aware of others in a way that infants are not aware of their parents: we are aware of others as having interests that make them potentially ‘unsympathetic’ to us, given our own interests, needs and desires. This awareness seems to militate against developing trust in others, where trust, as Dreyfus says, involves some confidence that others will act “sympathetically to our interests even when so doing might go against [their] own” (Dreyfus 2001, 70). So the question is how do we adult trusters reconcile the two: awareness of potential conflict and the risk of betrayal, on the one hand, with a capacity for genuine confident trust, on the other? I see two possibilities.

One possibility is to retreat from the awareness altogether, simply expecting in our trusting relationships the very stability and security we experienced in infancy—a kind of unconditional care that depends on erasing any threatening sense of difference between self and other. This is the hallmark of ‘security trust’. However, given the real existence of the other, how can this sense of total security be achieved except by erasing the other’s real identity, and superimposing on them a fantasized identity constructed in terms of just those needs and interests that resonate with our own? Although the allure of this strategy is clear, I think it’s also clear why it is inherently unstable and so highly vulnerable to breakdown. Moreover, since security-driven trusters fail to develop resources for negotiating difficulties when they arise, they are more liable than mature trusters to experience any breakdown of trust in terms of deep and traumatic betrayal—i.e. the sort of betrayal that is essentially irreparable.

A second possibility, clearly preferable to the first, is to reconcile our awareness of genuine difference with the capacity for genuine trust by giving up on the need for absolute security in our adult relationships of trust. Such security belongs to the days of our infancy. Nevertheless, what we can have in adult relationships of trust is potentially far more rewarding than mere security. Real difference invites the challenge of real discovery and the possibility of mutual enrichment. In trusting others who differ from ourselves, we create bonds of mutual concern, interest and support that provide a platform for exploring aspects of the world and our human condition that we might not get to on our own. Mature trust involves risk; but when all goes well, it makes other things possible that we would not, or could not, achieve by ourselves. The mature truster understands this dynamic and accepts what it implies, namely—*uncertainty*, some inevitable *divergence of interests*, and *potential conflict and breakdown*:

Uncertainty. In developing adult relationships of trust, mature trusters must deal with others whom they know only incompletely, partly because they have only limited evidence to go on, and partly because they recognize that individual character is not fixed in stone: like them, others are to some extent changeable, depending on their changing circumstances and experiences, and such changes can always affect on-going relationships of trust. Because of this incomplete knowledge, mature trusters accept that they must inevitably trust under conditions of uncertainty; trust means taking a calculated leap beyond the evidence.

Divergence of interests. No matter how much the truster’s interests seem to converge with trusted others, real difference means recognizing the inevitability of diverging interests.

Potential conflict and breakdown. The inevitability of diverging interests means problems will inevitably arise within relationships of trust. Mature trust involves recognizing this fact above all others. Functionally, it means developing the capacity to tolerate and negotiate differences when they arise. Consequently, mature trusters must develop resources to respond well when trust breaks down, despite the pain and disappointment such breakdowns may occasion.

Applying this analysis now to Julie's story, we might expect a mature truster to act rather differently on the discovery of her off-line identity occasion than an immature truster. For instance, as a friend of Julie's, the mature truster would seek to understand the reasons for Julie's deceit: Were her—or, rather, his—motivations inimical to friendship? Was he playing this character just for his own amusement, laughing at the women he was deceiving? Or did he have some possibly forgivable motive—say, a desire for intimacy that he felt he couldn't get by confessing his gender? (In fact, this was the explanation he gave.) Note that, in the context of mature trust, to judge that a motive is potentially forgivable does not entail the inevitable resumption of trust; but it does signal that, in the truster's view, there may be a way forward from such moments of betrayal to forging a new kind of relationship, premised on deeper mutual knowledge and understanding, and thereby paving the way for the possibility of renewed and better trust.

Mature trust is thus a kind of reasoned, explorative trust. Its primary impetus is not the need for security, but the desire to take calculated risks for the purpose of leading a richer human life. Mature trusters are individuals who trust, but who trust with care. They use care in three significant ways—namely, in *extending their trust*, in *monitoring their relationships of trust*, and in *responding well to others* when difficulties or breakdowns occur:

Extending trust. Mature trusters are discerning. They do not throw themselves into relationships of trust, but nor do they hold back when opportunities for trust present themselves. Because mature trusters are secure in themselves, they can afford to be exploratory in their trusting relationships, recognizing, but also testing, the kinds of conditions that effect relationships of trust in various ways. Thus, for instance, mature trusters are equipped to make distinctions between trust on the Internet and trust extended in other ways. Still, the fact that there are unique and interesting challenges for trust on the Internet does not stop the mature truster from facing those challenges and benefiting from the relationships made possible thereby.

Monitoring relationships of trust. For mature trusters, trust is not a one-off investment in a trusted other. It involves building and maintaining a relationship of trust that is sensitive to the changing attitudes and circumstances of truster and trustee.

Responding well to others when difficulties of breakdowns occur. Mature trusters understand that difficulties and breakdowns are part and parcel of relationships of adult trust. Despite the pain and disappointment inevitably caused by such occurrences, mature trusters work to understand the circumstances surrounding them. And if possible and desirable, they work to regenerate their

relationships of trust, perhaps even healing breaches in ways that make those relationships stronger than they were before.

In short, mature trusters are sensible about their trust, but they are also ready and willing to trust under conditions which allow them to balance the difficulties of engaging with fully autonomous and only partially known others against the many prospective rewards that come from relationships of trust.

Can the Internet provide such conditions? I don't see why not. Consequently, we have no good reason to insist that trust on the Internet of the sort I've been discussing cannot be rational. It can be rational just in case it provides mature trusters with the opportunities to engage well with others, developing the kind of responsive attentive relationships that we expect with our friends. It may still be somewhat limited trust—limited because it is uni-dimensional, confined to the medium of textual exchange. Relationships of trust are generally multimedia, deepening with the multifaceted ways we have of interacting with one another. But textual exchange is one of our best ways to bring mind into contact with mind, and heart eventually to heart, so I see no reason why Internet relationships cannot deepen over time in a similar kind of way. Of course, if they do so deepen, they are likely to spill over into life off-line.

3. The Presentation of Self in Internet Life

In the last section, I argued that trust on the Internet can be rational—that is to say, there is no reason why text-based interactions should not provide mature trusters with opportunities to engage well and responsibly with trusted others, deepening and broadening their relationships with these others over time. Of course, there are distinctive challenges to building trust on the Internet as we saw in Section 1, and mature trusters must proceed in reflective awareness of them. To review those challenges here:

The first has to do with the fact that Internet identity is relatively less grounded than identity off-line, allowing Internet personas to float somewhat free of the real world individuals who enact them. I called this the *Proteus factor*. One consequence of this is that individual identity on the Internet must be associated more directly with an agent's intersubjective properties—i.e., properties the agent manifests in relating to others—rather than, as is often the case off-line, with the agent's personal descriptive properties—i.e. properties, like age, sex, profession, appearance and so on, that pick the agent out as a particular identifiable person. Normally, who we take someone to be in this descriptive sense has a dramatic impact on how we expect them to act towards us, and so on how we act towards them in turn. This is the power of stereotyping. On the Internet, we are not quite deprived of this guide to another's likely profile of inter-subjective behaviour, since individuals will present themselves as possessing various personal descriptive properties. But since these descriptions are wholly within their control, they can be used more manipulatively than in life off-line to shape their recipients' intersubjective expectations. Julie's story is a prime example of such manipulation. How bad is this for Internet trust? As we

shall see in a moment, it may not be so bad after all. But it is something that individuals must be aware of if they are to trust well on the Internet.

The second challenge for Internet trust involves what I called the *Eliza factor*: Because Internet personas are manifested only in text, it allows others to project more freely onto them whatever fantasies they have about the off-line agents who enact the personas. This can have painful consequences. Because we will inevitably feel betrayed if the persons we trust do not live up to our expectations, it's important to have expectations that are realistically grounded if we are to trust reasonably and well. The problem of forming unrealistic expectations based on our fantasies about others are not unique to the Internet, but they can be greatly facilitated by the textual medium in which Internet exchanges are conducted.

While those who trust well may compensate for these liabilities of Internet interaction, we have seen that those who trust badly—immature or irrational trusters—may fall prey to them. Indeed, they may use the limitations of Internet interactions to avoid the responsibilities and risks of full adult trust. So, perhaps the best that can be said for trust on the Internet is that its special challenges do not rule out rational trust, but they certainly don't make it easy either. Thus, it may seem reasonable to conclude that there are no positive advantages to building relationships of trust on the Internet. But I think this judgement would be over-hasty. I close by considering some possible advantages of Internet trust that may outweigh the obvious liabilities and, in fact, are interestingly related to those liabilities.

So far my argument has been that the particular liabilities of Internet trust, which I've called the Proteus and Eliza factors, are not unique to the Internet; they are just greatly magnified by the Internet. Thus, problems of Internet trust make more salient problems that exist in any relationships of trust. For people present themselves to us in everyday life, masking some aspects of character, highlighting others; and on the basis of our limited exposure to them, we interpret the signs that they give us as favourable or unfavourable for trust. We can be manipulated in making these judgements by what they present to us; and we can also be manipulated in making these judgements by what we ourselves project onto them as a consequence of our needs and desires. Now, if building relationships of trust on the Internet makes more salient difficulties that are already present in forging any relationships of adult trust, then *by making these difficulties more salient* Internet use may force a practice of reflection on these difficulties and a direct engagement with them that engenders in turn more mature capacities for trust.

In support of this suggestion, I want to conclude by briefly considering a particular form of text-based multi-user interaction on the Internet that has become increasingly popular: the so-called MUD or virtual society. Unlike Internet conferences and chat rooms, MUDs involve a form of interaction in which the Proteus factor is, as it were, formalized. Individual players explicitly create and enact personas that are understood to be descriptively different from their off-line selves along any of a number of different dimensions: gender, ethnicity, age, profession, and even species (some MUD personas could at most be characterized

as ‘humanoid’, and humanoid only because of their interpersonal characteristics rather than their physical characteristics as described by those who enact them).

MUDs are extremely popular. No doubt for a variety of reasons, people are deeply attracted to enacting personas they make up themselves, and interacting with other such personas in a virtual world that is communally constructed. Are such activities pure escapism? Perhaps to a certain degree. But it is interesting to note how deeply invested dedicated MUD players become in this shared virtual world, in their relationships with other personas in that world, and most of all, in the personas they create for themselves. For instance, one such player remarked, “MUDS make me more what I really am. Off the MUD, I am not as much me” (Turkle 1996, 54). This statement seems to reflect a sentiment that is widely shared among dedicated MUD players. Presumably, it shows that they are able to lead lives in this virtual space and via their MUD personas that allow them to experience their own potential for creative activity and social engagement in more direct and rewarding ways than they can in ‘real life’. But why should this be?

I can think of two important reasons (which might be relevant to Julie’s case as well). The first involves avoiding, or at least modifying, the effects of stereotyping in the real world. As I pointed out earlier, people’s views about whom they are interacting with has a powerful effect on how they expect those individuals to behave and on how they should govern their behaviour towards them in turn. Stereotypes have a profound effect on the scope of individuals’ social agency. It may be impossible to escape these effects completely. But by playing with the personal properties of self-presentation, it may be possible to alter the shape of one’s social agency by provoking others to apply stereotypes that are different from those they would automatically apply in the real world. Thus, for example, many MUD players enact personas with the opposite gender from their own just to experience what is possible for them within a differently gendered social space (Bruckman 1996).

A second, perhaps more interesting reason involves self-development. Studies in social psychology show that when individuals publicly enact personality traits that they do not really take themselves to possess, this has a lasting effect on how they come to think of themselves—they think of themselves as now possessing those traits (Tice 1992). In other words, it seems that though the very process of having the traits they enact accepted by others as part of themselves, individuals themselves come to accept their behaviour as stemming more directly from who they really are. This may not be so surprising; but it does show that for many players, what may begin as a kind of enjoyable public make-believe becomes a powerful source of self-enablement and, hence, of self-development.

If these sorts of reasons explain why individuals report feeling ‘more like themselves’ in the context of role-playing games, I think it also clarifies that what matters most to agents’ sense of self-identity in the MUD is not who they are in terms of personal descriptive properties, but rather who they are in terms of their intersubjective properties—i.e., in terms of the traits they are enabled to enact in relation to other personas. And, of course, what matters to agents most in terms of the identities of other personas with whom they interact is the intersubjective

traits of these personas in turn. Relationships of trust and friendship in the MUD thus depend on the stability of players' intersubjective properties—hence, on their taking responsibility for maintaining their characters in relationships to one another, and of responding well to the concerns of one another when conflicts arise. Moreover, such responses have come to include, interestingly enough, explicit reflection on the difficulties of trust and civil behaviour in the MUD given the way players' identities are constructed and maintained. (There are special forums for such discussions on MUDs, usually in the form of a mailing list to all participants—e.g. **social-issues* in the MUD LambdaMOO).⁶

My point in discussing these virtual communities is just this: The fact that friendship and community survive, sometimes even flourish, in such forums shows the extent to which individuals can and do take responsibility for fomenting and maintaining relationships of trust, explicitly regulating their behaviour to compensate for conditions that destabilize and potentially undermine such trust. My claim is not that these relationships can substitute for relationships of trust in the real world, where the risks and vulnerabilities for self and others are presumably much more serious. My claim is rather that such conditions can make the difficulties inherent in relationships of trust more apparent to Internet trusters, thus challenging them to respond to these difficulties in ways that develop their own reflective capacities for engaging in relationships of mature trust both on and off-line. In the words of Sherry Turkle,

“Virtual personae can be a resource for self-reflection and self-transformation. Having literally written our on-line world into existence, we can use the communities we build inside our machines to improve the ones outside of them. Like the anthropologist returning home from a foreign culture, the voyager in virtuality can return to the real world better able to understand what about it is arbitrary and can be changed.” (Turkle 1996, 57)

Bibliography

- Baier, A. (1986), Trust and Antitrust, in: *Ethics* 96, 231–260
- Bechara, A./H. Damasio/D. Tranel/A. R. Damasio (1997), Deciding Advantageously before Knowing the Advantageous Strategy, in: *Science* 275, 1293–1294
- Bechara, A./D. Tranel/H. Damasio/A. R. Damasio (1996), Failure to Respond Autonomically to Anticipated Future Outcomes Following Damage to Prefrontal Cortex, in: *Cerebral Cortex* 6, 215–225
- Bruckman, A. S. (1996), Gender Swapping on the Internet, in: V. Vitanca (ed.), *CyberReader*, Needham Heights, 418–425
- Campbell, S. (1997), Love and Intentionality: Roxanne's Choice, in: R. Lamb (ed.), *Love Analyzed*, Boulder, 225–242
- Dreyfus, H. (2001), *On the Internet*, London-New York
- Hertzberg, L. (1989), On the Attitude of Trust, in: *Inquiry* 31, 307–22
- Holton, R. (1994), Deciding to Trust, Coming to Believe, in: *Australasian Journal of Philosophy* 72, 63–76

⁶ For further information on LambdaMOO, see <http://www.lambdamoo.info/>

- Jones, K. (1996), Trust as an Affective Attitude, in: *Ethics* 107, 4–25
- (2003), Emotional Rationality as Practical Rationality, in: C. Calhoun (ed.), *Setting the Moral Compass: Essays by Women Philosophers*, New York-Oxford
- McGeer, V. (2002), Developing Trust, in: *Philosophical Explorations* 5, 21–38
- Meltzoff, A./A. Gopnik (1993), The Role of Imitation in Understanding Persons and Developing a Theory of Mind, in: S. Baron-Cohen/H. Tager-Flusberg/D. J. Cohen (eds.), *Understanding Other Minds: Perspectives from Autism*, Oxford, 335–366
- Pettit, P. (1995), The Cunning of Trust, in: *Philosophy and Public Affairs* 24, 202–25
- Rostand, E. (1959), *Cyrano de Bergerac*, New York
- Stone, A. R. (1991), Will the Real Body Please Stand up?, in: M. Benedikt (ed.), *Cyberspace: First Steps*, Cambridge, 81–118
- Strawson, P. (1974), *Freedom and Resentment*, London
- Tice, D. M. (1992), Self-Concept Change and Self-Presentation: The Looking Glass Self is Also a Magnifying Glass, in: *Journal of Personality and Social Psychology* 63, 435–51
- Turkle, S. (1996), Virtuality and its Discontents, in: *American Prospect* 7, 50–57
- Walker, M. U. (forthcoming), *Fixing Responsibility: Essays in Moral Repair*, Cambridge
- Weizenbaum, J. (1976), *Computer Power and Human Reason*, San Francisco