

## SECURING TRUST ONLINE: WISDOM OR OXYMORON?

HELEN NISSENBAUM\*

INTRODUCTION.....	000
I. CONCEPTUAL AND TECHNICAL SCOPE.....	000
II. THE FUNCTION OF TRUST .....	000
III. CONDITIONS OF TRUST .....	000
A. <i>History and Reputation</i> .....	000
B. <i>Inferences Based on Personal Characteristics</i> .....	000
C. <i>Relationships: Mutuality and Reciprocity</i> .....	000
IV. THE SOLUTION: SECURITY.....	000
A. <i>Access Control</i> .....	000
B. <i>Transparency of Identity</i> .....	000
C. <i>Surveillance</i> .....	000
V. CAN TRUST BE SECURED?.....	000
A. <i>Securing Trust Versus Nourishing Trust</i> .....	000
B. <i>Security is no Panacea</i> .....	000
CONCLUSION: STRUGGLE FOR THE SOUL OF CYBERSPACE CONCLUSION.....	000

### INTRODUCTION

This essay is about trust in the online world.<sup>1</sup> It is not a manual on how to achieve trust, nor an empirical study of trust's presence or absence online. Rather, it is an attempt to show that the way we stipulate the conditions of the online world may be decisive for whether or not trust is achieved. Applying this perspective to a dominant vision of trust as security, the essay argues that if conditions are wrongly stipulated, then efforts to achieve trust may be misdirected, indeed, even thwarted.

The value of trust for a robust online world is obvious. Trust is a key to the

---

\* Member, School of Social Science, Institute for Advanced Study; Research Associate and Lecturer, University Center for Human Values, Princeton University.

<sup>1</sup> This paper is an outgrowth of a collaborative project with Edward Felten and Batya Friedman, and owes much to them. The work was supported by grants from the National Science Foundation, SBR-9729447 and SBR-9806234. I am enormously grateful to colleagues for probing questions and suggestions: Tamar Frankel, Jeroen van den Hoven, Rob Kling, and Mark Poster; for editorial and research assistance: Beth Kolko, Helen Moffett, Michael Cohen, and Hyeseung Song, Sayumi Takahashi, and Robert Young. Earlier versions of the paper were presented at: Computer Ethics: Philosophical Enquiry 2000, New York University School of Law: Conference on A Free Information Ecology, and Boston University School of Law: Conference on Trust Relationships.

promise the online world holds for great and diverse benefits to humanity—its potential to enhance community, enliven politics, hasten scientific discovery, energize commerce, and more. Trust in the layered infrastructures of hardware, software, commercial and institutional presences, and its people is a necessary condition for benefits to be realized. People shy away from territories they distrust; even when they are prepared to engage voluntarily, they stay only as briefly as possible. Without people, without participants, many of the visions will be both literally and figuratively empty. Trust would invigorate the online world; suspicion and insecurity would sap its vibrancy and vitality.

In exploring the issue of trust online, I turned to the work and insights of two communities of researchers, writers and practitioners. To learn about concerns relating specifically to trust *online*, I found much discussion “in the air” and also in an extensive literature spanning scholarly and trade publications, the popular media, government reports, and the Web itself. I found a second source of insights in the considerable works on trust by philosophers, social scientists and social theorists. These expanded, clarified and enriched the common sense conception of trust with which I began this exploration.

Animating the literature on the subject of trust *online* were two concerns. One, usually expressed by technical experts in security, was a concern over the fragility of technical systems. These experts worried that our vast networked information system—the network of networks including local private systems as well as public systems like the Internet, the Web, Cyberspace—is vulnerable to technical failure as well as malicious attack.<sup>2</sup> The second concern is over the success of e-commerce if consumers balk because they are fearful that they will be cheated, defrauded, have their credit card numbers stolen, or receive poor quality goods, or if businesses stay away fearing costly losses from such actions as failure to pay, repudiation of their commitments, and so on.<sup>3</sup>

---

<sup>2</sup> See, e.g., COMMISSION ON INFO. SYS. TRUSTWORTHINESS, NATIONAL RESEARCH COUNCIL, TRUST IN CYBERSPACE 1 (Fred B. Schneider, ed. 1999) (hereinafter TRUST IN CYBERSPACE) (“The widespread interconnection of networked information systems allows outages and disruptions to spread from one system to others; it enables attacks to be waged anonymously and from a safe distance . . .”)

<sup>3</sup> See, e.g., James P. Backhouse, *Security: The Achilles Heel of Electronic Commerce*, 35 SOC’Y 28, 28 (1998) (discussing security issues in e-commerce); Donna L. Hoffman et al., *Building Consumer Trust Online*, COMM. OF THE ACM Apr. 1999, at 80, 80 (addressing the trust issues between consumers and businesses in e-commerce); Robert Moskowitz, *Ask Yourself: In Whom Can You Really Trust?*, NETWORK COMPUTING 1, ¶ 1 (Jun. 15 1998) <<http://www.networkcomputing.com/911/911colmoskowitz.html>> (discussing the doubts that plague e-commerce); Pauline Ratnasingham, *Implicit Trust Levels in EDI Security*, 2 J. INTERNET SECURITY, 1, ¶ 1 (1999) <<http://www.addsecure.net/jisec/1999-02.htm>> (arguing that trust is an “important antecedent” for successful business relationships); Karl Salnoske, *Building Trust in Electronic Commerce*, 100 BUSINESS CREDIT 24, 24 (Jan. 1998) <<http://www.nacm.org/bcmag/bcarchives/1998/articles1998/jan/jan98art2.html>>

Although inspired by distinct sources, the proponents of these two concerns converge in their vision of the likely shape of a solution; namely, a particular suite of technical security mechanisms that they believe will induce users to trust these networked information systems. Through strong mechanisms of security and security oriented practices, we should seek to create “trusted,” or rather, trustworthy<sup>4</sup> systems that would, in turn, induce consumers to trust providers of goods and services, providers to trust consumers, and in general, engender a climate of trust online.<sup>5</sup> So conspicuous has been the vision of trust through security portrayed by these two groups that it currently occupies the mainstream—in part because there are no equally persistent, competing interpretations, and in part, because talk of trust online is relatively new and the mainstream view relatively uncontested. Later in this paper, I shall say more about these mechanisms, but here I would like to label this common vision with a slogan: trustworthiness as security, or trust through security.<sup>6</sup>

This essay is an evaluation of the vision of trust through security. Its thesis, guided by conceptions of trust developed in the theoretical and empirical work of social scientists and philosophers, is that the online landscape thus

---

(commenting that both businesses and consumers regard transaction security as their biggest concern); Dennis D. Steinauer et al., *Trust and Traceability in Electronic Commerce*, 5 STANDARD VIEW 118, 118 (1997) (exploring “technology or other processes that can help increase the level of confidence . . . in electronic commerce”); David Woolford, *Electronic Commerce: It’s All a Matter of Trust*, 25 COMPUTING CANADA 18, ¶ 1 (May 7, 1999) <<http://www.plesman.com/Archives/cc/1999/May/2518/cc251813b.html>> (arguing that electronic deals suffer from the problems of “authenticity and integrity”).

<sup>4</sup> A misuse of language persists within the technical computer security community: proponents of a particular security device invariably use the term “trusted” to signal their faith that the system in question is trustworthy. This usage is misleading, as it suggests a general acceptance of the device in question when in fact it is the duty of the proponents to argue or prove that it is indeed worthy of this acceptance.

<sup>5</sup> See, e.g., Alfaraz Abdul-Rahman & Sphen Hailes, *A Distributed Trust Model*, in. NEW SECURITY PARADIGMS WORKSHOP 48, 48 (1998) (discussing the weaknesses of current security approaches for managing trust); DEPARTMENT OF DEFENSE TRUSTED COMPUTER SYSTEM EVALUATION CRITERIA (visited July 1, 1999) <<http://www.all.net/books/orange>> (classifying computer systems into four divisions of enhanced security protection); Rohit Khare & Adam Rifkin, *Weaving a Web of Trust* (visited Jan. 13, 2001) <<http://www.w3journal.com/7/s3.rifkin.wrap.html>> (1997) (“develop[ing] a taxonomy for how trust assertions can be specified, justified and validated”); Michael K. Reiter, *Distributing Trust with the Rampart Toolkit*, COMM. OF THE ACM, Apr. 1996, at 71, 71 (describing group communication protocols which distributes trust among a group).

<sup>6</sup> Although I do not discuss their work here, I must acknowledge another community of researchers: namely those interested in computer-human interaction, who are concerned with ways to elicit trust through the design of user interfaces. See, e.g., Ben Schneiderman, *Designing Trust into Online Experiences*, COMM. OF THE ACM, Dec. 2000, at 57, 58-59 (outlining certain steps, such as disclosing patterns of past performance and enforcing privacy and security policies, that designers can take to encourage trust in online relationships).

envisioned will not be conducive to trust and trustworthiness; trust online will not be achieved through security because that vision is founded on a misconstrued notion of trust, missing the point of why we care about trust and making mistaken assumptions about human nature along the way. Before attending to this central thesis, we must first set realistic boundaries for the scope of this essay. Because the technological realm of which we speak is so extensive and intricate, and the conceptual domain of trust so broad and varied, we must make some qualifications and simplifying assumptions.

### I. CONCEPTUAL AND TECHNICAL SCOPE

In its broadest sense, the online world we speak of could cover the entire technological system, vast and powerful, that sits at the hub of almost all other parts of the critical infrastructures of society, controlling, and in some cases conjoining energy, commerce, finance, transportation, education, communication, and more, and as such, affecting almost all modes of social, community, cultural and political life.<sup>7</sup> This essay does not address the system as a whole—the vast and powerful grid that connects and controls satellites, nuclear devices, energy, the stock exchange, and so forth. Instead, it focuses on those parts of the system directly experienced by the ordinary people, who in increasing numbers, use it to talk, conduct business transactions, work, seek information, play games, and transact with public and private institutions. At present, this means the World Wide Web (the Web) and the various servers (computers), conjoined networks, people, and institutions that comprise it. It means the realm that at times interacts with the realities of the offline world, and at other times, fragments into an apparently independent and separate reality that some writers and participants have taken to calling Cyberspace, or the “virtual” world.

Neither does this essay cover everything that *trust* could mean. Trust is an extraordinarily rich concept covering a variety of relationships, conjoining a variety of objects. One can trust (or distrust) persons, institutions, governments, information, deities, physical things, systems, and more. Here, I am concerned with two ways that trust is used. One is as a term describing a relationship between one person (a trustor) and another (the trustee). Although, in practice, the trustee position could be filled by almost anything, here I limit consideration to cases where the trustee is a being to which we are willing to attribute intentions, motivations, interests, or reasons, and might also refer to as “agents.” Central to this category, I would place people—individually and in groups; I would also be willing to include organizations,

---

<sup>7</sup> See TRUST IN CYBERSPACE, *supra* note 2, at 12-23 (evaluating whether and to what degree we can rely on existing networked information systems that support our critical infrastructures). This report urged a set of actions to increase trustworthiness and limit our vulnerability to harm, even catastrophe, that might result from failures due to malfunction or malicious attack. See *id.* at 240-55 (outlining the Commission’s conclusions and recommendations).

communities, and institutions. However, I exclude from my discussion at least one quite common reference to trust in the online context: trust in the networked, digital information systems themselves, in the layered hardware and software that individually comprise the micro-systems and the macro-system that is formed by these. This is not because of any deep-seated disagreement with those who write about trust in relation to networked information systems or information and communications technology and worry about the dependability of these systems, their resilience to various forms of failure and attack, and their capacity to protect the integrity of online interactions and transactions. My reasons are pragmatic. These cases are sufficiently distinct from one another that they deserve separate (but equal) treatment. Following others, I use the term “confidence” to refer to trust in systems, recognizing that trust in the online world begins with confidence in systems, but does not end there.<sup>8</sup>

## II. THE FUNCTION OF TRUST

Why we care about trust online is a question that provokes several lines of response. One is to seek an explanation for why the online realm appears especially problematic, that is, why we worry particularly about trust *online*. We shall return to this later, but first, let us consider why we might care about trust at all. I treat this daunting question in a more limited way by thinking about trust’s function in order to generate a sense of the ways in which trust contributes in positive ways to our lives. Even in this limited sense, I can give here only a compressed and selective account, merely sampling from the extensive literature on the general subject of trust and its value, allowing my account to be shaped by the particular focus of this paper.

We might quickly agree with the general view that trust is good, though a more considered reaction is likely to yield the view that trust is good for certain ends. It is an instrumental good whose ends, in our common experience, are usually good, but need not be. Scholarship endorses this qualified position on trust as a phenomenon that is implicated in the achievement of many valued aspects and institutions of individual and social life.<sup>9</sup> This work has revealed the benefits of trust for individuals (both for those who trust as well as those who are trusted) for relationships and for communities. In the case of individuals, there are the psychological benefits both of being trusted and of trusting, of not being stricken with paranoia and suspicion. Clearly there is a lot to say about all these benefits, but I would particularly like to draw attention to one aspect of the value of trust for individuals observed by Niklas Luhmann, a social theorist whose profound

---

<sup>8</sup> See ADAM B. SELIGMAN, *THE PROBLEM OF TRUST* 19 (1997) (arguing that trust in systems entails confidence in a set of institutions).

<sup>9</sup> See, e.g., Annette Baier, *Trust and Antitrust*, 96 *ETHICS* 231, 232 (1986) (“There are immoral as well as moral trust relationships, and trust-busting can be a morally proper goal.”)

work on trust has been widely influential.

Luhmann characterizes trust as a mechanism that reduces complexity and enables people to cope with the high levels of uncertainty and complexity of contemporary life.<sup>10</sup> Trust makes uncertainty and complexity tolerable because it enables us to focus on only a few possible alternatives.<sup>11</sup> Humans, if faced with a full range of alternatives, if forced to acknowledge and calculate all possible outcomes of all possible decision nodes, would freeze in uncertainty and indecision. In this state, we might never be able to act in situations that call for action and decisiveness. In trusting, Luhmann says, “one engages in an action as though there were only certain possibilities in the future.”<sup>12</sup> Trust also enables “co-operative action and individual but coordinated action: trust, by the reduction of complexity, discloses possibilities for action which would have remained improbable and unattractive without trust—which would not, in other words, have been pursued.”<sup>13</sup> According to this account, trust expands people’s capacity to relate successfully to a world whose complexity, in reality, is far greater than we are capable of taking in.

Trust’s rewards extend beyond the individual, leavening many important relationships. Some, like the relationships between friends, lovers, siblings, husbands and wives, parents and children, mentors and students, are predicated on trust. But even in impersonal and formal relationships, trust plays a critical role: for trade and commercial transactions, for relationships between professionals (caregivers, healers, lawyers, etc.) and their clients, between employers and employees, between constituents and their political representatives.<sup>14</sup>

The possibilities for action increase proportionately to the increase in trust—trust in one’s own self-presentation and in other people’s interpretation of it. When such trust has been established, new ways of behaving become possible; jokes, unconventional initiatives, bluntness, verbal short cuts, well-timed silences, the choice of delicate subjects, etc. When trust is tested and proven in this way, it can be accumulated by way

---

<sup>10</sup> NIKLAS LUHMANN, *Trust: A Mechanism for the Reduction of Social Complexity*, in TRUST AND POWER: TWO WORKS BY NIKLAS LUHMANN 8 (photo. reprint 1988) (1979) (“[T]rust constitutes a more effective form of complexity reduction.”)

<sup>11</sup> See *id.* at 20 (noting that trust evolves from past experiences that can guide future actions).

<sup>12</sup> *Id.* at 20.

<sup>13</sup> *Id.* at 25.

<sup>14</sup> See generally FRANCIS FUKUYAMA, TRUST: THE SOCIAL VIRTUES AND THE CREATION OF PROSPERITY 7 (1995) (illustrating examples for the need for trust in economic life); Baier, *supra* note 9, at 239 (commenting that ordinary individuals must trust the mailman and the plumber to do their jobs properly); Lawrence C. Becker, *Trust as Noncognitive Security about Motives*, 107 ETHICS 43, 51 (1996) (discussing trust of government officials); Russell Hardin, *Trustworthiness*, 107 ETHICS 26, 33 (1996) (explaining how economic institutions are trustworthy with their customers); Philip Pettit, *The Cunning of Trust*, 24 PHIL. & PUB. AFF. 202, 204-05 (1995) (discussing the trust placed in a city bus driver).

of capital.<sup>15</sup>

This idea of trust as capital—social capital—has been developed and popularized by Robert Putnam in his study of Italian communities and later work suggesting a decline in social capital in American society.<sup>16</sup> With each trust-affirming action, trust accrues in communities as capital, to stabilize, to exert control, and to induce cohesion and solidarity, to be there to tap in troubled times. The value of trust in social and associational life, not necessarily mediated through social capital, is something that other political philosophers have endorsed. Philip Pettit, for example, stresses the strength and solidarity that trust can engender, concluding, like Putnam “that trust is a precious if fragile commodity in social and political life”;<sup>17</sup> it is characteristic of flourishing civil societies.<sup>18</sup> Trust among citizens may be the magic ingredient that helps undergird political and civil stability in multicultural societies;<sup>19</sup> trust is an “important lubricant of a social system;”<sup>20</sup> it is the basis for modern solidarity.<sup>21</sup> Trust by individuals of such institutionalized authority as government may sustain a citizenry’s engagement in a social system, and may even stave off highly volatile and disruptive reactions one might normally expect in the wake of harms that citizens believe to have been caused by these authorities.<sup>22</sup>

From these and other works we learn that trust is especially important in complex, varied, and somewhat unpredictable, personal, social and political contexts where much is at stake. Trust facilitates cooperation and success within civil and political society; it enriches individuals’ lives by encouraging activity, boldness, adventure, and creativity, and by enriching the scope of individuals’ relationships with others. It is not surprising, therefore, that an interest in trust should grow just as the realm we known as Cyberspace, the

---

<sup>15</sup> LUHMANN, *supra* note 10, at 40 (footnote omitted).

<sup>16</sup> ROBERT D. PUTNAM, *MAKING DEMOCRACY WORK: CIVIC TRADITIONS IN MODERN ITALY* 169 (1993) (arguing that social capital increases when it is used and diminishes when it is not used).

<sup>17</sup> Pettit, *supra* note 14, at 225.

<sup>18</sup> *See id.* at 202 (arguing that society where people trust one another will most likely function “more harmoniously and fruitfully” than society devoid of trust); *see also* FUKUYAMA, *supra* note 14, at 47 (arguing that “sociability is critical to economic life because virtually all economic activity is carried out by groups rather than individuals”); PUTNAM, *supra* note 16, at 170 (asserting that trust is an “essential component” of social capital).

<sup>19</sup> *See* Daniel M. Weinstock, *Building Trust in Divided Societies*, 7 *POL. PHIL.* 263, 263-83 (1999).

<sup>20</sup> KENNETH J. ARROW, *THE LIMITS OF ORGANIZATION* 23 (1974).

<sup>21</sup> *See* SELIGMAN, *supra* note 8, at 73 (arguing that solidarity must include some element of trust).

<sup>22</sup> *See* Becker, *supra* note 12, at 51 (voicing that the majority of U.S. citizens trust the motives of public officials enough to combat the effect of receiving negative information about them).

Internet, the Web, the Global Information Infrastructure, burgeons, just as it is beset by deep and difficult questions about authority and governance,<sup>23</sup> just as it crosses a threshold of complexity where participants, in increasing numbers, turn to the online world for many of the experiences, relationships, community-life, information and commercial interactions that once they lived entirely in so-called “real” space.

Before turning attention to the online world, though, I note two qualifications. As an instrumental good, trust may, on occasion, serve evil ends. Trust between partners in crime increases the chances of criminal success, and social solidarity among oppressive communities strengthens their efficacy.<sup>24</sup> The same might be said about other instrumental values, even privacy and freedom, which we see as overwhelmingly positive even as we seek for ways to limit their exercise for the sake of evil ends. A second qualification is that trust is not appropriate for every situation and relationship. In negotiating with a used-car salesman or with a sworn enemy, which may both be necessary, we prefer strategies other than trust. In choosing a bank or bakery, for example, trust also may not be crucial.

Returning to the online world, we would expect that trust here holds a key to similar good ends: improving the quality of personal experiences, relationships, and communal and civic life, and stabilizing governance. We can expect that more people and institutions will “buy in” to the online world, will engage with others online, if there is sufficient trust. If a climate of trust can be established on the Net, if attitudes of trust toward partners in electronically mediated transactions can be achieved, then the online world will thrive, it will attract information, it will be lively with interaction, transaction and association. This will attract further investment of all kinds, which in turn will fuel participation, and so on. Conversely, if people do not trust interactions mediated electronically, they will minimize them; they will be cautious and suspicious in their dealings, they will not place information and creative works on the web, they will not engage in E-commerce, they will not indulge in MUDS, MOOS, E-lists, B-boards, Listservs, chatrooms, buddy lists, electronic banking, and more. A great resource will be wasted.

---

<sup>23</sup> Consider, for example, controversies over governance of a range of issues from speech and gambling to the allocation of Domain Names. For one of the classic (and controversial) positions on Internet governance, see David R. Johnson & David Post, *Law and Borders—The Rise of Law in Cyberspace*, 48 STAN. L. REV. 1367, 1367 (1996) (arguing that Cyberspace requires different laws than the laws that govern geographically-defined territories).

<sup>24</sup> For a similar critique of social capital, see Alejandro Portes & Patricia Landolt, *Social Capital: Promise and Pitfalls of its Role in Development*, 32 J. LAT. AM. STUD. 529, 546 (2000) (commenting that “one must not be over-optimistic about what enforceable trust . . . can accomplish”).



### III. CONDITIONS OF TRUST

At the same time that proponents of the Internet acknowledge the role of trust as a key to flourishing activity, interaction and institutional growth, they recognize certain distinctive features of the online realm that may interfere with building and sustaining trust. To see how these features may interfere, let us think first about conditions that generally have been associated with the formation of trust. We initiate this inquiry by asking about the mechanisms that govern trust. To what factors are tendencies to trust (or not to trust) systematically responsive? What factors influence the tendency to trust other people, groups, and institutions? One may sensibly ask these questions whether one holds that trust is a species of belief (or expectation) or that it is a non-cognitive attitude, a matter of some disagreement among theorists and social scientists. Those, like Baier, who assert a version of the former view, ask about reasons that may systematically undergird trust, and may even subject it to judgments of rationality or irrationality.<sup>25</sup> Those, like Becker, who defend a non-cognitive account of trust, can nevertheless agree that trust is systematically responsive to a variety of factors.<sup>26</sup> For purposes of this paper, it will not be necessary to settle the question, as long as our agnosticism does not prevent us from drawing on empirical and analytic work that links trust with a variety of phenomena that function systematically as its cues, clues, or triggers, whether these function as reasons or merely as causes.

Admittedly, the factors listed below reflect my interest in the relevant distinctiveness of the online context and should not be taken as a complete theory of the causes of trust. My efforts may also rub against views on trust, like Adam Seligman's, which would reserve the concept of trust for an even more qualified sub-category of attitudes than the one I have articulated. Seligman, for example, prefers to use a term like confidence for cases where similarity, or roles and other structured relationships, induce a positive expectation with respect to the other.<sup>27</sup> To engage further on this point of disagreement—interesting as it is—would deflect us too far from the main subject of the paper. It is important, though, to acknowledge the difference between my more ample and Seligman's more austere concepts. One way to reconcile the difference would be to suggest that followers of Seligman's usage recast the concern of this paper as being one about trust, faith, confidence, and familiarity online.

#### A. *History and Reputation*

One of the most convincing forms of evidence that others merit trust is their

---

<sup>25</sup> See Baier, *supra* note 9, at 259 (arguing that in some instances it is more prudent to distrust, rather than to trust).

<sup>26</sup> See Becker, *supra* note 12, at 58 (noting that a "proper sense of security is a balance of cognitive control and noncognitive stability").

<sup>27</sup> See SELIGMAN, *supra* note 8, at 16-21 (explaining the difference between trust and confidence).

past behavior. If they have behaved well in the past, protected our interests, have not cheated or betrayed us, and, in general, have acted in a trustworthy manner, they are likely to elicit trust in the future. If they have disappointed in the past, then we will tend not to trust them. Where we have not built a history of direct interaction with others, we may refer to the experiences of others; that is to say, we may be influenced by their reputations.

B. *Inferences Based on Personal Characteristics*

A trusting attitude may be triggered by the presence of perceived qualities in the other. Philip Pettit identifies four: virtue, loyalty, prudence<sup>28</sup> and a desire for the good opinion of others,<sup>29</sup> all qualities that influence whether a person will trust those who are seen to have them. Pettit writes, “[t]o be loyal or virtuous or even prudent is, in an obvious sense of the term, to be trustworthy. It is to be reliable under trust and to be reliable, in particular, because of possessing a desirable trait.”<sup>30</sup> The fourth quality, namely, a desire for the good opinion of others, although less deserving of our admiration, is nevertheless a powerful mechanism for preventing betrayals of trust.<sup>31</sup> Accordingly, Pettit recommends against calling the person who chases good reputation *trustworthy*, preferring a more modest commendation of *trust-responsiveness*, or *trust-reliant*.<sup>32</sup> Though not in direct disagreement with Pettit’s characterization, Adam Seligman offers a different perspective, drawing attention to the importance of familiarity, similarity and shared values as triggers of trusting attitudes.<sup>33</sup> What we know about someone, what we may infer on the basis of “their clothing, behavior, general demeanor,”<sup>34</sup> may lead us to judgments about their values and moral commitments, especially telling if we judge them to be similar to ours. A common religious background, high school, neighborhood, or traumatic experience (e.g., having fought in the same war), affects how confident we are in predicting what others will do and how inclined we are to rely on them. Though related to loyalty, these considerations are not identical. When one depends on a loyal cousin, for example, one counts on the family relationship to induce trust-reliance in one’s cousin. Where trust is triggered by familiarity and, perhaps, a perception of

---

<sup>28</sup> Pettit, *supra* note 14, at 210 (arguing that the mechanisms of trust can explain why “trust builds on trust”).

<sup>29</sup> *See id.* at 203 (commenting that many are not proud of this trait).

<sup>30</sup> *Id.* at 211.

<sup>31</sup> *See id.* at 203 (arguing that people regard their desire for the good opinion of others as a disposition that is hard to shed).

<sup>32</sup> *See id.* at 207 (arguing that “[w]here trust of this kind materializes and survives, people will take that as a token of proof of their being well disposed toward one another, so that the success of the trust should prove to be fruitful in other regards”).

<sup>33</sup> SELIGMAN, *supra* note 8, at 69 (arguing that familiarity relates to the “human bond” rooted in identity).

<sup>34</sup> *Id.* at 69.

shared values, a trustee does not necessarily count on these qualities to cause trustworthy behavior; the trustee merely forms expectations regarding the likely actions of these others.

#### C. *Relationships: Mutuality and Reciprocity*

Aside from personal qualities, the relationship in which one stands to another may bear on the formation of trust. The presence of common ends can stimulate trust. Such cases of mutual ends occur when a person is “in the same boat” as another. When I fly in an airplane, for example, I place trust in the pilot partly because he is in the plane with me and I presume that we have common, or confluent, ends; our fates are entwined for the few hours during which we fly together.

Reciprocity is slightly different, but it, too, can be grounds for trust. In a reciprocal relationship, we trust others not because we have common ends, but because each of us holds the fate of others in our hands in a manner of tit-for-tat. This may occur, for example, when people are taking turns. The agent whose turn is first deals fairly, reliably, or responsibly with the other because soon the tables will be turned. The relationship of reciprocity admits of great variability. In some cases, there is clear and imminent reversal of roles (this year I am chair of our department, next year you take over); in others it is more generalized (I might donate money to the Cancer Foundation hoping that when I become ill, these funds will somehow help me). Reciprocity is evident in communities that are blessed with a climate of trust, helping those in need and trusting that when they are in need, others will help them.<sup>35</sup>

#### D. *Role Fulfillment*

There is another, perhaps more compelling reason for trusting the pilot of my airplane. After all, the pilot would not trust me, in spite of our common interest in staying alive. Crucial to my trusting the pilot is that he is a pilot, and being a pilot within the framework of a familiar system has well-articulated meaning. I know what pilots are supposed to do, I am aware of the rigorous training they undergo, the stringent requirements for accreditation, and the status of airlines within a larger social, political and legal system. Several of the authors already mentioned have discussed the importance of roles to the formation of trust.<sup>36</sup>

---

<sup>35</sup> See PUTNAM, *supra* note 16, at 172 (arguing that reciprocity undergirds social trust, which facilitates cooperation in communities).

<sup>36</sup> See, e.g., SELIGMAN, *supra* note 8, at 22 (arguing that the concept of social role has been “fundamental to modern sociological analysis”); Baier, *supra* note 9 at 256 (arguing that people trust others to perform their roles in society); Pettit, *supra* note 14, at 221 (arguing that divisions among people in a community are likely to reduce the chances of people from different sides trusting one another).

### E. *Contextual Factors*

One of the most intriguing factors to affect our readiness to trust, beyond those that are tied to what we know about the other, is the nature of the setting in which we act.<sup>37</sup> Such settings can be construed quite locally as families, communities, and towns, or can extend to ones as large and diffuse as nations and countries.

Four elements seem relevant. The first is publicity: a setting in which betrayal and fidelity are routinely publicized is likely to be more conducive to trust-reliance, and consequently trust, than a setting in which people can effectively hide their deeds—especially their misdeeds. The second is reward and punishment: settings in which rewards and sanctions follow trustworthiness and betrayal respectively, are likely to induce trustworthiness and trust. Thirdly, where reward and punishment for fidelity and betrayal are not systematically available, promulgation of norms through other means can effectively shape behaviors, and establish a climate of one sort or another. What norms are conveyed through parables, education, local lore, songs, fables, and local appraisal structures? Do they condemn betrayal and celebrate fidelity or do they mock gullible marks of confidence tricks, and disdain cuckolded spouses while proffering admiration to the perpetrators?<sup>38</sup> Finally, a society can nurture a trusting climate by setting in place, through public policy or other means, various forms of “trust insurance,” to provide safety nets for those whose trust is betrayed.<sup>39</sup> A simple example of such a policy is the current arrangement of liability for credit card fraud, which must surely increase people’s willingness to engage in credit transactions.

## IV. OBSTACLES TO TRUST ONLINE

Knowing a little more about trust, we return to trust online. To begin, we observe that the online world is relatively new. Novelty, or unfamiliarity, can in itself stall the formation of trust. Beyond sheer novelty, however, there are more specific features of the online world that bear on the formation and sustenance of trust, which cloak many of the aspects of character and personality, nature of relationship, and setting that normally function as

---

<sup>37</sup> See LUHMANN, *supra* note 10, at 78-85 (discussing the conditions necessary for trust to be formed); Russell Hardin, *The Street-Level Epistemology of Trust*, 21 POL. & SOC’Y 505, 514 (1993) (asserting that the “terrible vision of a permanent underclass in American city ghettos may have its grounding in the lesson that the children of the ghetto are taught . . . that they cannot trust others”); Pettit, *supra* note 14, at 222 (arguing that a society in which trust is found only in small family groups might become very cynical); Weinstock, *supra* note 19, at 263-83.

<sup>38</sup> See LUHMANN, *supra* note 10, at 84 (commenting on how “complex and richly varied the social conditions for the formation of trust are”).

<sup>39</sup> See Hardin, *supra* note 37, at 522 (discussing social mechanisms that generate trust); Pettit, *supra* note 14, at 220 (arguing that the “trust-responsiveness mechanism” has implications for institutional design); Weinstock, *supra* note 19, at 263-83.

triggers of trust or as reasons for deciding to trust (or distrust).

#### A. *Missing Identities*<sup>40</sup>

In its current design, the medium allows agents to cloak or obscure identity. In many of their online transactions, agents are not compelled to relinquish the identities of their off-line selves. Although this ability to engage online anonymously is beneficial in a number of ways, it shrinks the range of cues that can act as triggers of trust or upon which people base decisions to trust. If we imagine identity as a thread upon which we string the history of interactions with others, then without that thread we lose the ability to learn from to past experiences of either vindicated trust or betrayal. Lacking information about sustained identity means we are also deprived of the means of learning from the experiences of others whether an agent is trust reliant, as the construction of reputation is hampered—if not precluded altogether—where identity is unknown.

Lacking knowledge of an agent's sustained identity also means that we may not have the information necessary to recognize the nature of the sustained relationships in which we stand, for example, whether it is reciprocal or cooperative. Finally, because identity is also bound up with accountability, people might presume that anonymous agents are less likely to act responsibly. As a result, they would be less inclined to trust.

#### B. *Missing Personal Characteristics*

There is an opacity not only with respect to others' identities, but with respect to many of the personal characteristics that affect (heighten or diminish) attitudes of trust. Online, we are separated from others in time and space; we lack cues that may give evidence of similarity, familiarity, or shared value systems. We may not know the other's gender (male, female, or "other"), age, race, socioeconomic status, occupation, mode of dress, or geographic origins. We lack the bodily signals of face-to-face interaction. Are we communicating with a 14-year-old girl or a 57-year-old man posing as a 14-year-old girl? Are we selling a priceless painting to an adolescent boy or to a reputable art dealer?<sup>41</sup> Are we sharing a virtual room with an intriguing avatar or a virtual rapist?<sup>42</sup> We must conduct transactions and depend on

---

<sup>40</sup> There is far more complexity to this issue than I need, or am able, to show here. See, e.g., Helen Nissenbaum, *The Meaning of Anonymity in an Information Age*, 15 INFO. SOC'Y 141, 141 (1999) (discussing anonymity and what it means to protect it); Kathleen Wallace, *Anonymity*, 1 ETHICS AND INFO. TECH. 23, 23 (1999) (offering a definition of anonymity).

<sup>41</sup> In 1999, a thirteen-year-old boy from Haddonfield N.J. participated in Ebay auctions, bidding away \$3.2 million dollars on items like a van Gogh sketch and a 1971 Corvette convertible. His parents were successful in freeing themselves from responsibility for these transactions. See *Boy Bids \$3M at Online Site*, AP ONLINE (Haddonfield), Apr. 30, 1999, available in 1999 WL 17062405 (reporting the exploits of the eighth-grade online bidder).

<sup>42</sup> See Julian Dibbell, *A Rape in Cyberspace; or, How an Evil Clown, a Haitian Trickster*

others who are separated not only by distance but also by time, who are disembodied in many of the ways that typically contribute to our sense of their trustworthiness.

### C. *Inscrutable Contexts*

The novelty and difference of the online environment lead not only to the veiling of properties that affect the formation of trust: the settings themselves are frequently inscrutable in ways that affect readiness or inclination to trust. One casualty is role definition, because, at least for now, we cannot rely on traditional mechanisms for articulating and supporting social, professional and other roles. Even with roles that appear equivalent to offline counterparts, for example, “shopkeeper,” we lack the explicit frameworks of assurances that support them. For the roles that have emerged in cyberspace (like “sysops,” avatars, bulletin board moderators, and so on) that do not have obvious counterparts offline, their duties and responsibilities are even less defined and understood.

Just as roles are still relatively unformulated, so are such background constraints and social norms regarding qualities like fidelity, virtue, loyalty, guile, duplicity, and trickery. Are we sure that betrayal will be checked, that safety nets exist to limit the scope of hurts and harms, and so on? Although there is evidence of various groups—social groups, interest groups, cultural groups—vying for domination of their norms, the territory remains relatively uncharted, further compounded by the global nature of the medium. Participants, especially the majority, who are not strongly identified with any one of these groups, can rightly be confused. For them, the most rational stance may be one of caution and reserve.

It is important to note that what I call inscrutability of contexts has a double edge. Many people have observed that it is precisely this quality of Cyberspace that is so liberating, enticing, promising. Enthusiasts invite you to participate *because* it is new, different, better, seamless, immediate, unstuffy, truly democratic, and so forth. I am not sure, therefore, that even if we could, the immediate solution to the problem of inscrutability is a wholesale transfer of existing norms.

## IV. THE SOLUTION: SECURITY

Given that we are deprived of the usual cues and triggers, what steps can we take to sustain trust? A cohort of security experts, security-minded systems managers, government oversight bodies, and proponents of e-commerce advocate developing a particular suite of security mechanisms. This suite of mechanisms would assure freedom from harm for online agents by allowing them to act and transact online in safety. The idea is that safety will build

---

*Spirit, Two Wizards, and a Cast of Dozens Turned a Database into a Society*, in *FLAME WARS: THE DISCOURSE OF CYBERCULTURE* 237, 237-40 (Mark Dery, ed. 1994) (describing a fictional virtual rape in an online multi-user domain).

trust.

Computer security is not a new concern, but rather has developed alongside the development of computing itself, responding to changes in the technology and the needs of its rapidly expanding range of applications. What is relatively new, however, is the close linking of the purposes of security with the idea of trust. There are, no doubt, a number of reasons why trust has entered the picture as one of the values that guides security. There are the historical and sociological reasons mentioned above; there is the ubiquity and power of the technical infrastructure; there is the fundamentally open (or insecure) architecture of the technical infrastructure. In this paper, we focus on the part of the picture that links security mechanisms to trust via the missing clues, cues and triggers. Or, rather, we explore the security mechanisms developed in the name of trust—those that function, in part, to restore those missing clues, cues and triggers.

What follows is a brief overview of some of the efforts in answer to this call, which I have organized according to three categories: 1) Access Control; 2) Transparency of Identity; and 3) Surveillance. The categories, which are largely my own construction, are an obvious simplification of the broad range of work in computer and network security. My intent is not to describe categories explicitly adopted by computer security experts themselves, nor to suggest that there is a monolithic effort of people and projects, but to provide explanatory clarity relevant to the purposes of discussing trust. The categories reflect functionality and not underlying structural similarities, and, as we shall soon see, are highly interrelated.

#### A. *Access Control*

One of the earliest worries of computer security, from the time when computers were stand-alone calculators and repositories of information, has been to guard against unwanted access to the computer and its stored information, to maintain the integrity of the information, and to control distribution of the valuable and limited resource of computational power. Early on, the security mechanisms developed to prevent illegitimate and damaging access would involve anything from passwords to locked doors.<sup>43</sup> The demands on computer security mechanisms expanded and became more complicated as networks and interactivity evolved. Vulnerability to intrusion grew because networks opened new means of infiltration: email, file transfer, and remote access—which could not be stemmed by locked doors. The infamous Morris Worm, which received widespread national attention, jolted all users into noticing what security experts must certainly have feared: that it was merely a matter of time before vulnerabilities in theory would be

---

<sup>43</sup> This is what I mean by organizing according to functionality. Structurally, a password is a very different device to a locked door, but in relation to this aspect of computer security, namely access control, they are effectively the same.

exploited, in practice.<sup>44</sup>

The Internet, and in particular the Web, has further expanded the modes and extent of interactivity while, at the same time, exposing participants to new forms of unwanted access and attack. The old fears remain: namely, infiltration by unauthorized persons (hackers, crackers, etc.), damage to information and systems, disruptive software flowing across the Net, information “stolen” as it traverses the networks, terrorists and criminals invading the infrastructure and bringing down critical systems. And new fears emerge: “evil” websites that harm unsuspecting visitors, web-links diverted from intended destinations to others, and disruptive applets—mini-applications that visitors to Websites can download onto their own systems to enable them to enjoy more extensive services from that site. Khare and Rifkin note, “[w]hile doing nothing more serious than surfing to some random Web page, your browser might take the opportunity to download, install, and execute objects and scripts from unknown sources.”<sup>45</sup> To view a video clip, for example, visitors might need to download a player program in addition to the video files themselves; they may download a mini-spreadsheet program to view and interact with financial information provided by a financial services company. In the process of downloading the appropriate application, however, the user’s computer system is infected with a harmful, often devastating, applet. Greater interactivity spells greater vulnerability and a need for more extensive protections. Bruce Schneier, a computer security expert, comments on the level of vulnerability after one particular round of attack-and-repair:

Looking back to the future, 1999 will have been a pivotal year for malicious software: viruses, worms, and Trojan horses (collectively known as “malware”). It’s not more malware; we’ve already seen thousands. It’s not Internet malware; we’ve seen those before, too [sic]. But this is the first year we’ve seen malware that uses e-mail to propagate over the Internet and tunnel through firewalls. And it’s a really big deal.

What’s new in 1999 is e-mail propagation of malware. These programs—the Melissa virus and its variants, Worm.ExploreZip worm and its inevitable variants, etc.—arrive via e-mail and use e-mail features in modern software to replicate themselves across the network. They mail themselves to people known to the infected host, enticing the recipients to open or run them. They don’t propagate over weeks and months; they propagate in seconds. Anti-viral software cannot possibly keep up. . . .

One problem is the permissive nature of the Internet and the computers

---

<sup>44</sup> See Ashley Dunn, *Computer World Battles Faster-Moving Viruses Technology*, LA TIMES, Oct. 4, 1999, at C1, C7 (reflecting on the “notorious” outbreak of the Morris Worm and explaining that an internet security clearinghouse was created in response to the damage done by the Morris Worm).

<sup>45</sup> Khare & Rifkin, *supra* note 5, at ¶ 4.



attached to it. As long as a program has the ability to do anything on the computer it is running on, malware will be incredibly dangerous.

And anti-virus software can't help much. If a virus can infect 1.2 million computers (one estimate of Melissa infections) in the hours before a fix is released, that's a lot of damage. . . .

It's impossible to push the problem off onto users with "do you trust this message/macro/application" messages . . . . Users can't make good security decisions under ideal conditions; they don't stand a chance against a virus capable of social engineering. . . .

What we're seeing here is the convergence of several problems: the permissiveness of networks, interconnections between applications on modern operating systems, e-mail as a vector to tunnel through network defenses as a means to spread extremely rapidly, and the traditional naivete of users. Simple patches won't fix this. . . . A large distributed system that communicates at the speed of light is going to have to accept the reality of viral infections at the speed of light. Unless security is designed into the system from the bottom up, we're constantly going to be fighting a holding action.<sup>46</sup>

Working within the constraints of current network and system architectures, security experts have developed a toolkit of mechanisms to protect people and systems against unwanted and dangerous access. One reason why demands on such a toolkit are considerable is because the agents of unwanted access may be bits of code, like applets, and not only people. Standard techniques like passwords remain in use, fortified where needed by such mechanisms as "firewalls," which are software barriers built around systems in order to make them impermeable except to people or code that is "authorized."<sup>47</sup> Cryptographic techniques are used to protect the integrity and privacy of information stored in computers; such techniques also protect against theft and manipulation as information travels across networks. Some protection is offered against treacherous applets—for example, one that might reformat a user's hard drive, or leak private information to the world—through security features built into JAVA that limit what applets can do. There are, however, regular announcements of flaws in this security.<sup>48</sup> There is fundamentally no

---

<sup>46</sup> Bruce Schneier, *Risks of E-mail Borne Viruses, Worms, and Trojan Horses*, 20 RISKS DIGEST 2, ¶¶ 1, 6, 9, 11, 12, 13 (June 17, 1999) <<http://catless.ncl.ac.uk/Risks/20.45.html>>.

<sup>47</sup> See, e.g., COMMITTEE ON INFORMATION SYSTEMS TRUSTWORTHINESS, NATIONAL RESEARCH COUNCIL, TRUST IN CYBERSPACE 134-37 (Fred B. Schneider ed., 1999) (defining firewalls and identifying them as one of the mechanisms used to prevented unwanted access to computer systems).

<sup>48</sup> See, e.g., Gary McGraw & Edward Felten, *Understanding the Keys to Java Security*, JAVAWORLD, May 1, 1997, available in 1997 WL 28334788 (reporting that a "code-signing hole" had been found in Java software); Richard King, *Java Sun's Language is Scoring Some Early Points with Operators and Suppliers*, TELE.COM, May 1, 1996, available in

known technical means of differentiating “good” from “bad” applets. How could there be except in some possible future when computers would be able discern categories of human values?

### B. *Transparency of Identity*

The people and institutions of the online world have diverse tastes when it comes to identification. Some are happy to link themselves to their full-blown offline identities, while others prefer to maintain independent virtual selves. Among the second group, some are happy to maintain consistent identities represented by “handles” or pseudonyms, while others prefer full anonymity. The goal of security efforts in this category is to give more transparent access to online agents in order to stave off at least some of the threats and worries that follow from not knowing with whom one is dealing. Identifiability is considered particularly useful for recognizing malevolent or mischievous agents. And in general, it helps answer some of the questions that trust inspires us to ask: is there a recognizable and persistent identity to the institutions and individuals behind the myriad of websites one might visit? Can we count on agents online to keep their promises? For the sake of e-commerce, how do we prevent malicious agents from posing as legitimate customers or service providers, and conducting bogus transactions, tricking and defrauding legitimate participants? In other words, we strive to reintroduce identifying information, at least as much as is needed to create a history, establish a reputation, hold agents accountable, and so on.

Security efforts have focussed on the task of making identity sufficiently transparent to protect against these and other betrayals and harms in an effort to build what Lawrence Lessig has called “architectures of identification.”<sup>49</sup> Mostly, they are interested in developing a strong link between a virtual agent and a physical person through a constellation of information that is commonly seen as proving identity even offline.<sup>50</sup> Security experts are investigating the promise of biometric identification—for example through fingerprints, DNA profiles, and retinal images. Cryptographic techniques are deployed to authenticate users, computers, and sources of information by means of digital signatures and digital certificates working within a socially constructed system of Certification Authorities, trusted third parties who vouch for the binding of cryptographic keys to particular identities—persons and institutions. These

---

1996 WL 16760663 (noting that “[s]everal security flaws have been reported since Sun [Microsystems, Inc.] announced Java”).

<sup>49</sup> See LAWRENCE LESSIG, *CODE AND OTHER LAWS OF CYBERSPACE* 34-35 (1999) (identifying three common architectures of identity used on Internet as passwords, “cookies,” and digital certificates).

<sup>50</sup> *But cf.* Helen Nissenbaum, *The Meaning of Anonymity in an Information Age*, 15 INFO. SOC’Y 141, 143 (1999) (arguing that the information age’s capacity to aggregate and analyze the data necessary to identify an individual, even without access to a name, presents a new challenge to protecting anonymity, where society desires to do so).

same mechanisms are intended to prevent repudiation by agents of commitments or promises they may have made.

Schemes of identification, even the attenuated forms, work hand-in-hand with access control, because controlling access almost never means preventing everyone from using a system or the information in a system. It almost always means distinguishing the sanctioned, legitimate users from the illegitimate. In the case of applets, because direct examination can provide only imperfect evidence, we may rely on what is known about who sent them for another source of discrimination between “good” and “bad” applets.<sup>51</sup> “Trust management systems” are offered as integrated mechanisms for identifying and authenticating the identity of those people, information, and code that affect us, and are also supposed to authenticate an applet’s origins. The Snow White fairytale offers an irresistible comparison: if Snow White had known the true identity of the bearer of the apple, she could have avoided the fateful bite.

Security experts seem always to be engaged in a Sisyphusian battle, warding off attack, repairing system flaws, closing up loopholes and “backdoors,” and devising new layers of protection; a process that ends, temporarily at least, until the next attack occurs. Outspoken security experts accept that this is an inevitable consequence of the “open” architecture of the Internet and Web, which many consider to be fundamentally insecure.<sup>52</sup> As a result, we live with an unstable equilibrium of relative comfort until the latest, more devastating intrusion is made public; there is a flurry of reaction, followed by relative comfort, and so on.

### C. Surveillance

A third layer overlaid upon the security offered through access control and transparency of identity is surveillance: we keep an eye on things both in order to prevent harms and also to apprehend perpetrators after harm has been done. Surveillance can involve active watching and tracking, which can be fairly fine-grained, as demonstrated by the monitoring software that many business organizations have installed on their computer systems. Or they can be relatively coarse-grained, as are some “intrusion detection” systems, where real-time monitoring issues an alarm in response to suspicious or unusual activity to be further investigated if necessary.<sup>53</sup> Surveillance can also involve

---

<sup>51</sup> Microsoft Explorer’s security is based on this principle.

<sup>52</sup> See Schneier, *supra* note 46, at ¶ 9 (“One problem is the permissive nature of the Internet.”)

<sup>53</sup> This seems to be the form of the Federal Intrusion Detection Network (FIDNet) system proposed by the National Security Council and endorsed by the Clinton administration to protect government computers. See Marc Lacey, *Clinton Outlines Plan and Money to Tighten Computer Security*, N.Y. TIMES, Jan. 8, 2000, at A14 (identifying FIDNet as part of the Clinton Administration’s larger computer security plan); see also *White House Fact Sheet: Cyber Security Budget Initiatives*, U.S. NEWSWIRE, Feb. 15, 2000, available in 2000 WL 4141378 (outlining the Clinton Administration’s budget initiatives

passive recording (reifying) of digital trails. Popular means include logging and auditing, which creates records of activity through which authorities can sift at a later time. Logging and auditing helped authorities identify David Smith as the creator of the Melissa virus.<sup>54</sup>

#### V. CAN TRUST BE SECURED?

The claim we must examine is that through an array of mechanisms, such as firewalls, biometrics, digital signatures, intrusion detection, auditing, and so forth, trust will be secured online. We might adjust the claim slightly to apply to an idealized world in which we have perfected these mechanisms of access control, establishing reliable markers of identity and maintaining a watchful eye through surveillance. Will we thus secure trust?

Given the framing of the problem of trust online that we have seen so far, the claim has prima facie plausibility because the mechanisms in question appear to be a means of restoring some of the triggers of trust that elude us online. Strong and smart walls, and limits on the flow of information and range of interactivity establish “safe” zones; greater transparency of identity through authentication allows participants to steer clear of “suspicious” agents. By exposing identities or, at least, crucial dimensions of identities, agents—individuals, organizations, and computers—may more effectively make judgments about trustworthiness, and decide whether others are “safe bets”. Mechanisms of non-repudiation restore accountability. This, then, is the compelling current generated by the proponents of security and e-commerce.

In spite of its prima facie plausibility, however, I will argue that security, or

---

related to cyber security for fiscal year 2001); *see generally* THE WHITE HOUSE, DEFENDING AMERICA'S CYBERSPACE: NATIONAL PLAN FOR INFORMATION SYSTEMS PROTECTION, VERSION 1.0: AN INVITATION TO DIALOGUE, EXECUTIVE SUMMARY 15 (2000) (discussing various government intrusion detection systems). The FIDNet proposal has met with significant opposition from various civil liberties groups. *See, e.g.*, John Markoff, *The Strength of the Internet Proves to Be Its Weakness*, N.Y. TIMES, Feb. 10, 2000, at C1 (noting that FIDNet caused alarm among civil libertarians who said it would be used to curtail privacy on the internet); *see also* Patrick Thibodeau, *Senate Hears Objections to Cyberalarm*, COMPUTERWORLD, Feb. 7, 2000, at 25, available in LEXIS, News Library, U.S. News, Combined File (reporting on privacy group's testimony before the U.S. Senate Judiciary Subcommittee on Technology, Terrorism and Government Information).

<sup>54</sup> *See* John Leyden, *Melissa's Creator Faces 'Hard Time'*, NETWORK NEWS, Apr. 14, 1999, at 7, available in LEXIS, News Library, U.S. News, Computing & Technology file (reporting that America Online assisted federal and state law enforcement agents in identifying David Smith as the creator of the Melissa virus); Lee Copeland, *Virus Creator Fesses Up—Admits to Originating and Disseminating Melissa*, COMPUTER RESELLER NEWS, Sep. 6, 1999, available in LEXIS, News Library, Newspaper Stories, Combined Papers (noting that America Online tracked Smith down by tracing the virus to a list server in New Jersey); Hiawatha Bray, *N.J. man charged in computer virus case*, THE BOSTON GLOBE, Apr. 3, 1999, at A1, available in LEXIS, News Library, Newspaper Stories, Combined Papers (noting that America Online assisted the government agents in identifying Smith).

rather, the particular vision of security occupying the mainstream will not, as promised, bring about trust. I argue this not because I think security is unimportant, but because the ends of trust online are not well served by this mainstream vision of security. The rhetoric is misguided because when the proponents of security and e-commerce would bind trust too closely to security, they threaten to usurp a concept as rich and complex, as intensely social, cultural and moral as trust, for merely one slim part of it. The mistake is not merely semantic; it has a weighty practical edge. Pursuing trust online by pursuing the complete fulfillment of the three goals of security would no more achieve trust and trustworthiness, online—in their full-blown senses—than prison bars, surveillance cameras, airport X-ray conveyer belts, body frisks, and padlocks, could achieve offline. This is so because the very ends envisioned by the proponents of security and e-commerce are contrary to core meanings and mechanisms of trust.

There are two ways in which security misses the mark: it overshoots trust and it undershoots it.

#### A. *Securing Trust Versus Nourishing Trust*

Let us begin with the first critique, namely, that security as commonly prescribed may actually quash trust. Here, an excursion back to theoretical and empirical studies of trust is useful. Trust, we learn, is an attitude. It is almost always a relational attitude involving at least a trustor and a trustee. In this relation of trust, those who trust accept their vulnerability to those in whom they place trust. They realize that those they trust may exercise their power to harm, disappoint, or betray; yet at the same time, they regard those others “as if” they mean well, or, at least, mean no harm. Trust, then, is a form of confidence in another, confidence that the other, despite a capacity to do harm, will do the right thing in relation to the trustor. For the philosopher Annette Baier, trust is “accepted vulnerability to another’s possible but not expected ill will (or lack of good will) toward one;”<sup>55</sup> trust is the “reliance on others’ competence and willingness to look after, rather than harm, things one cares about which are entrusted to their care.”<sup>56</sup> For Russell Hardin, “trust involves giving discretion to another to affect one’s interests.”<sup>57</sup> In a similar vein, Adam Seligman holds trust to be “some sort of belief in the goodwill of the other, given the opaqueness of other’s intentions and calculations.”<sup>58</sup> Francis Fukuyama adds a social dimension to his account, describing trust as the “expectation that arises within a community of regular, honest, and cooperative behavior, based on commonly shared norms, on the part of other members of that community.”<sup>59</sup>

---

<sup>55</sup> Baier, *supra* note 9, at 235.

<sup>56</sup> *Id.* at 259.

<sup>57</sup> Hardin, *supra* note 37, at 507.

<sup>58</sup> SELIGMAN, *supra* note 8, at 43.

<sup>59</sup> FUKUYAMA, *supra* note 14, at 26.

Usually trust involves more than the trustor and trustee; there is almost always an object with respect to which the trustor trusts the trustee.<sup>60</sup> For Annette Baier, this is demonstrated in her example of trusting the plumber to take care of the pipes in her home, but not to take care of her daughter; trusting a babysitter to take care of her daughter, but not to take care of the pipes in her home.<sup>61</sup> A person might entrust even her life to a friend, but not her heart. In the online world, there is similar discretion over not only whom one is prepared to trust, but with what one is prepared to entrust them; for example, many consumers have learned that they can trust Amazon Books to deliver their orders, but not trust them with their personal information.<sup>62</sup>

Most relevant to our concern here, however, is a theme common to all the works that I have studied, namely the essential connection between trust and vulnerability. When people trust, they expose themselves to risk. Although trust may be based on something—past experience, the nature of one's relationships, etc.—it involves no guarantees. As Hardin writes, trust is “inherently subject to the risk that the other will abuse the power of discretion.”<sup>63</sup> In trusting, we are acknowledging the other as a free agent, and this is part of the exhilaration both of trusting and being trusted. Where people are guaranteed safety, where they are protected from harm via assurances—if the other person acted under coercion, for example—trust is redundant; it is unnecessary. What we have is certainty, security, and safety—not trust. The evidence, the signs, the cues and clues that ground the formation, that give evidence of the reasonableness of trust must always fall short of certainty; trust is an attitude without guarantees, without a complete warranty.<sup>64</sup> When we

---

<sup>60</sup> See Baier, *supra* note 9, at 236 (analyzing trust as a relationship in which “A trusts B with valued thing C,” and in which B is given discretionary powers with respect to C); Hardin, *supra* note 36, at 506 (“To say ‘I trust you’ seems almost always to be elliptical, as though we can assume some such phrase as ‘to do X’ or ‘in matters Y.’”); Weinstock, *supra* note 19, at 263-83.

<sup>61</sup> Baier, *supra* note 9, at 245 (“We take it for granted that people will perform their role-related duties and trust any individual worker to look after whatever her job requires her to. The very existence of that job, as a standard occupation, creates a climate of some trust in those with that job.”)

<sup>62</sup> See Goldberg et al., *Trust Ethics and Privacy*, 81 B. U. L. REV. \_\_, \_\_ nn. 70 & 71 (2001) (discussing changes to Amazon.com's privacy agreement, protecting customer information, that resulted in reduced protections). [ES—this is a symposium article; the note numbers are estimates based on status of editing as of 2/5/01. The footnotes in the Goldberg article are citations to websites, including Amazon's, showing the changes to the privacy agreement]

<sup>63</sup> Hardin, *supra* note 37, at 507.

<sup>64</sup> See LUHMANN, *supra* note 10, at 20 (noting that trust is based in part on familiarity, history, and past experiences); *id.* at 24 (arguing that trust always involves the risk that the harm resulting from a breach of trust may be greater than the benefit to be gained by trusting); Pettit, *supra* note 14, at 208 (arguing that irrespective of how one defines risk-taking, trust always involves putting oneself in a position of vulnerability whereby it is

constrain variables in ways that make things certain—i.e. safe—we are usurping trust’s function. Trust is squeezed out of the picture.

No loss, some, like Richard Posner, would say: “[b]ut trust, rather than being something valued for itself and therefore missed where full information makes it unnecessary, is, I should think, merely an imperfect substitute for information.”<sup>65</sup> According to Posner’s position, if we must choose between trust—and consequently, vulnerability—on the one hand, and certainty, on the other, then certainty must win.

In practice, however, such a choice has significant consequences, which are as evident online as off. In a world that is complex and rich, the price of safety and certainty is limitation. Online, we do not have the means at our disposal of assuring safety and certainty without paying this price: a streamlining and constraining of the scope and nature of online interactions, relationships, and community; a limiting of the range and nature of allowed activity and scope and nature of interaction; a need to make a priori judgments about with whom we will or will not interact; and an acceptance of greater transparency and surveillance.<sup>66</sup> In general, then, we trade freedom and range of opportunity for this certainty and safety.

The link between trust and vulnerability seems both to be conceptual and empirical. The conceptual claim is that whatever the feeling or attitude one experiences when acting and anticipating in a context of certainty and safety, it cannot be trust; this is not what trust means. The empirical conjecture, which has occupied the work of several scholars, is that in a context of complete certainty, the material conditions needed to induce and nourish trust are absent.<sup>67</sup> Trust does not flourish in a perfectly secure environment for reasons that are very different than the reasons trust does not flourish in a hostile, threatening environment. For trust to develop between an individual and either another individual or an organization, the trustor must somehow have had the opportunity to test the other agent and have had them pass the test. Luhmann explains the crucial role of uncertainty in the process of building trust:

First of all there has to be some cause for displaying trust. There has to be defined some situation in which the person trusting is dependent on his

---

possible for the other person to do harm to the trustor); Weinstock, *supra* note 19, at 263-83.

<sup>65</sup> Richard Posner, *The Right of Privacy*, 12 GA. L. REV. 393, 408 (1978).

<sup>66</sup> There has been discussion in the media about the Clinton administration’s proposals to monitor both governmental and private networks for signs of terrorist and criminal activity. See, e.g., Robert O’Harrow, *Computer Security Proposal Is Revised: Critics Had Raised Online Privacy Fears*, WASH. POST, September 22, 1999, at A31, available in LEXIS, News Library, Newspaper Stories, Combined Papers (reporting that civil liberties groups welcomed changes to the Clinton Administration’s original proposals, in particular limitations on automatic data collection); see also *supra* note 53 (discussing the Clinton Administration’s proposal for, and reaction to, enhanced computer network security programs).

<sup>67</sup> See LUHMANN, *supra* note 10, at 15 (noting that “trust increases the ‘tolerance of uncertainty,’” and explaining that “mastery of events” (i.e. knowledge) can replace trust).

partner; otherwise the problem does not arise. His behaviour [sic] must then commit him to this situation and make him run the risk of his trust being betrayed. In other words he must invest in what we called earlier a 'risky investment.' One fundamental condition is that it must be possible for the partner to abuse the trust . . . .<sup>68</sup>

When we are placed in a context in which we depend on others for our well being and are assured, guaranteed by whatever means, that these others are prevented and restrained and therefore incapable of harming us, then the context, though safe and secure, is not one that nourishes trust. No test has been given; none has been passed. The variables that theorists and empirical scientists have identified as trust-inducing may signal the reasonableness of trust in a particular setting, but when grounds are transformed into guarantees of good behavior, trust disappears, replaced not by distrust, but perhaps by certainty. In the presence of a violent psychopath whose limbs are shackled, one feels not trust, but, at best, safety.

Another empirical basis for doubting the efficacy of security to deliver trust is that boxing people in is a notoriously bad strategy for inducing trustworthiness or even trust-reliance. Constraining freedom directly or indirectly through, say, surveillance may backfire and have the opposite effect. Roderick Kramer, in reviewing empirical work in the social sciences, notes:

Ironically, there is increasing evidence that such systems can actually undermine trust and may even elicit the very behaviors they are intended to suppress or eliminate. In a recent discussion of this evidence, Cialdini (1996) identified several reasons why monitoring and surveillance can diminish trust within an organization. First, there is evidence that when people think their behavior is under the control of extrinsic motivators, intrinsic motivation may be reduced (Enzle & Anderson 1993). Thus, surveillance may undermine individuals' motivation to engage in the very behaviors such monitoring is intended to induce or ensure.<sup>69</sup>

Philip Pettit's observations reinforce this result:

[C]ertain intrusive forms of regulation can be counter-productive and can reduce the level of performance in the very area they are supposed to affect . . . . If heavy regulation is capable of eradicating overtures of trust, and of driving out opportunities for trusting relationships, then it is capable of doing great harm.<sup>70</sup>

The many inducements at the disposal of individuals and institutions to encourage trustworthiness are most effective when they operate indirectly. Above all, people need to perceive a choice. By means of these inducements, including sanctions and rewards, clearly articulated norms, education,

---

<sup>68</sup> *Id.* at 42.

<sup>69</sup> Roderick M. Kramer, *Trust and Distrust in Organizations: Emerging Perspectives, Enduring Questions*, 50 ANN. REV. PSYCHOL. 569, 591 (1999).

<sup>70</sup> Pettit, *supra* note 14, at 225.



character development, and so on we may increase the incidence of trust as well as trust-reliance. On the other hand, if we go too far, and deny the possibility of choice, we deny what is fundamental to trusting relationships and climates of trust. Symbols of trust can be exhibited in small but clear ways, as illustrated at the service counter of a popular downtown, café. A discreet sign says, "At our busy times, please be respectful of those waiting for tables." We trust, we do not coerce; we cannot assure decency, but we offer patrons an opportunity to demonstrate their trustworthiness.

#### B. *Security is no Panacea*

If the earlier criticism was that security, following the trajectory of development described above, overshoots the mark and creates an environment that does not allow trust to take root and flourish, then this alternative criticism that security does not go far enough. For even though security mechanisms promise to reduce our vulnerability in some ways, they leave us vulnerable in other ways that are relevant to the prospect of trust online. This loophole is all the more worrisome because, having achieved some modes of safety through security, we might fail to notice its significance until considerable damage is done.

To clarify, it will be useful to set in place a simplification, framing what is at stake in terms of "insiders" and "outsiders." Experts in computer security are worried about outsiders: malicious, avaricious, incompetent, or simply unauthorized outsiders who may break into our online space, damage or steal information, and destroy or compromise our systems. They develop security mechanisms to keep outsiders where they belong—outside—and to help spot or identify outsiders in order to take appropriate—preventative or punitive—action.

Far less systematic attention is paid to the threat of insiders, those agents—individuals and organizations—who, by degrees, have sanctioned access to our space. These agents, who count among the respectable, socially sanctioned, reputable members of online society, engage in actions that many citizens of the online world dislike, resent, or even consider harmful. They track our Web activities, they collect and use personal information without our permission, they plant "cookies" on our hard drives, they hijack our browsers while they download ads, they fill our mailboxes with spam, and they engage in relentless commercialism. Some of these insiders—perhaps not the "respectable" ones—"troll" our discussion groups, afflict us with hateful, inflammatory, mean-spirited emails ("flame"), send us threatening chain mail, and even attack our virtual selves.<sup>71</sup> In other words, even if the walls of security keep outsiders outside, they do not curtail the agents and activities that, behind the veil of respectability and legal sanction—sometime ambiguity—make online citizens skittish, cautious, and resentful. Such security barriers do not address various

---

<sup>71</sup> See Dibbell, *supra* note 42, at 239-42 (describing a fictional virtual rape in an online multi-user domain).

forms of activity that are fully capable of engendering a climate of suspicion and distrust online even if we are successful in our projects to secure the online world.

Even in the physical world, attention only to the threats of outsiders misses a universe of possibility linked closely with the presence of trust and trustworthiness. Consider the more familiar case of physical safety. To protect ourselves from bodily harm, many of us go to great lengths: we stay clear of dangerous parts of town, we affix padlocks to our doors and install burglar alarms in our homes, and we support the use of Closed Circuit Television (CCTV) in public spaces. Homicide statistics, however, tell a curious story: when the relationship of the killer to victim is known, we find that only twenty two percent of killers are strangers—the proverbial outsiders.<sup>72</sup> Others are spouses, friends, and acquaintances. Betrayal comes from those who are allowed within our spheres of safety, within our safe zones.

My intention is not to launch into paranoid realms of suspicion and universal distrust. It is to illustrate that keeping outsiders out need not assure safety. A wall of defense against malicious outsiders does not defend against the threats posed by sanctioned insiders, who energetically defend their “right” to exercise online freedoms—by means of cookies, misleading registrations, matching, mining, and so on. They are, arguably, chipping away at trust just as surely as amoral hackers are. As much as the latter, they are capable of causing a dangerous ebb in the abundant social capital we currently enjoy in life online.

Because it is in the nature of trust to be conservative—both to ebb and to grow—the results of these transgressions may not be immediately evident.<sup>73</sup> That the transgressions I speak of are capable of undermining trust, however, is implied by several of the works that have shaped this essay. One example is found in a long-term study of e-commerce, which shows that consumers’ trust is related to their understanding of how information about them is treated; it wanes if they think that it will not be held in confidence.<sup>74</sup>

Another important insight that explains why interventions like the familiar

---

<sup>72</sup> See U.S. DEP’T OF JUST., BUREAU OF JUSTICE STATISTICS: SELECTED FINDINGS, VIOLENT CRIME 3 (1994) (reporting that “in murders where the relationship between the victim and the offender was known, 44% of the victims were killed by an acquaintance, 22% by a stranger, and 20% by a family member”).

<sup>73</sup> See Becker, *supra* note 12, at 50 (noting that “ordinary life” provides substantial anecdotal evidence that most people have personal relationships in which they remain “trustful despite the known untrustworthiness of others”); cf. Paul Slovic, *Perceived Risk, Trust and Democracy*, 13 RISK ANALYSIS 675, 677 (1993) (describing trust as fragile and identifying “the asymmetry principle,” by which trust is usually created slowly, but destroyed in an instant, often by a single event).

<sup>74</sup> See Hoffman et al., *supra* note 3, at 82 (concluding that the primary barriers to consumers providing demographic data to Web sites are related to trust and noting that over 72% of Web users indicated they would provide demographic data if the Web sites would provide information regarding how the collected data would be used).

suite of security mechanisms cannot fully induce trust is that trust is as sensitive, if not more so, to motives and intentions as it is to actions and outcomes. It is in the goodwill of the other, Lawrence Becker has argued, that we trust or fail to trust, not necessarily their actions.<sup>75</sup> As long as we believe that others are well intentioned towards us, our trusting attitude towards them will survive a great deal of bad news: “incompetence, mendacity, greed, and so forth.”<sup>76</sup> This holds for the relation of citizens to government as well as among persons. According to Becker, only when citizens begin to attribute the poor performance of governments to deviant motivations—e.g., corruption or inappropriate power seeking—will they “respond in ways that are . . . volatile and disruptive.”<sup>77</sup> Citizens’ trust, it seems, is able to survive incompetence, at least for a while. In a similar vein, Paul Slovic, an expert on risk assessment, reports that the extent to which citizens are willing to accept societal risk due to technological innovation is related to their degree of confidence in the motives of those in charge.<sup>78</sup>

Similar ideas emerge in Tom Tyler’s research on public trust of police and the courts. Tyler is interested in variables that affect citizens’ confidence in legal authorities, their readiness to accept outcomes, and their evaluation of the quality of decision-making and fairness of procedures.<sup>79</sup> He finds that the most important variable is trust in the motives of authorities.<sup>80</sup> This Tyler calls motive based trust: “Motive based trust is distinct from judgments about whether or not authorities behave as anticipated. It involves an inference about the ‘spirit’ or ‘motive’ that will shapes [sic] behavior, not what specific behavior will occur.”<sup>81</sup> One of Tyler’s somewhat surprising findings is that in

---

<sup>75</sup> See Becker, *supra* note 12, at 59 (arguing that a person’s loss of confidence in another person’s motivations does more harm to the relationship than when the other person proves to be “merely unreliable or not credible”).

<sup>76</sup> *Id.* at 51.

<sup>77</sup> *Id.* at 59.

<sup>78</sup> See Slovic, *supra* note 73, at 680 (contrasting French and American citizens’ reaction to risks associated with nuclear power and noting that the French public’s acceptance of the risks is partly related to the public trust in the state-run nuclear program, which has a reputation for emphasizing public service over profits).

<sup>79</sup> See Tom Tyler, *Trust and Law Abidingness: A Proactive Model of Social Regulation*, 81 B. U. L. REV. \_\_ (2001) (advocating a “proactive model of social regulation that is based upon encouraging and maintaining public trust in the character and motives of legal authorities”). [the parenthetical quote is from page 2 of the draft]

<sup>80</sup> See *id.* at \_\_ (“Motive based trust is central to situations in which people rely upon fiduciary authorities.”) [the parenthetical is quoted from page 8 of the draft]; see also *id.* at \_\_ (summarizing results of an empirical study and concluding that trust is an important factor in shaping people’s reactions to their experience with legal authorities because one, “people who trust the motives of the authority with whom they are dealing are more willing to defer to that authority;” and two, “trust leads to more positive feelings about the legal authority involved”). [second parenthetical is from page 20 of the draft]

<sup>81</sup> *Id.* at \_\_. [Citation is to page 9 of the draft. Note that the draft version in the carrel

brushes with law enforcement and legal authorities, people's positive reactions are tied more strongly to inferred motives than even to whether or not the outcomes of their cases were favorable to them.<sup>82</sup>

The significance of these ideas for the purposes of this section is to emphasize that the behavior of many sanctioned, established, powerful individuals and organizations is capable of undermining trust even when the actions they undertake, such as web-tracking, for example, are not immediately aggressive or harmful. In these cases, when we learn of such activities we may find them ambiguous. What would matter to us for purposes of trust are the motivations behind the behaviors. As long as we are not able to read people's minds, directly assessing motives and intentions is difficult, often impossible. So we usually find ourselves drawing on as many indirect sources as possible, sometimes resorting to subtle detection and artfulness.

One important indirect source of others' intentions is their interests. When, for example, a politician seeking office expresses concern for a particular situation, voters might attribute the expression not to genuine feeling, but to an interest in being elected. In a case of this type, as much as we welcome and praise the action, it may not serve as grounds for trust as long as we see it emanating from a motive of vote seeking. In the case of web tracking—and more generally, information gathering and commercialism—we might initially be willing to read positive meaning into such practices. As time goes by, and we take measure of the distance between our own interests and those of the trackers (profit and potency), we begin to reinterpret those same actions as forms of betrayal. Actions that at first seem neutral or even friendly can come to be seen as sinister when interpreted in light of inferred motives and intentions.

We all need to interact, even cooperate, with others whose interests are not consistent with our own and may even conflict with ours. In such cases, we transact cautiously, ever on the lookout for betrayal, sometimes seeking protections from the most egregious harms, betrayals and exploitation. So trust in such cases remains elusive.

If we choose not to pursue policies for the online world that aim to contain the pursuit of avaricious interests that are contrary to those of the citizens of the Net, we are, I fear, planting the seeds of general distrust. People may continue to participate in this arena, but will do so with caution and a sense of

---

uses "shapes" rather than "shape," but this is probably a drafting error, which should be fixed in the editing process]

<sup>82</sup> See *id.* at \_\_ ("In the context of a specific personal experience with a legal authority, people are willing to voluntarily defer based upon their belief that the authorities are acting in a trustworthy manner. They infer trustworthiness from the justice of the actions of the authorities.") [parenthetical is from page 48 of the draft]; see also *id.* at \_\_ (discussing the opportunities police officers and judges have to developing public good will by justifying outcomes by reference to the public's moral values, in the outcome context, and treating people fairly in the procedural context). [second parenthetical is from page 50 of the draft article.]

wariness, wisely so, in interactions with those whose interests run contrary to our own, and whose actions may be annoying, bothersome, intrusive, or even threatening. Guardedness will be the norm.

Those who would pursue security in the name of trust do us this disservice. They focus on the outsider, the aberrant individual or organization, the trickster, the evil hacker, and the scam artist. These are the villains from which security would protect us. In proportion to actual harm done to individuals online, too much attention is paid to the aberrant individual, the trickster, and the evil hackers lurking outside the borders of civilized society online. The media play up dramatic cases: the Melissa virus, spies who infiltrate systems and sell secrets to our enemies, or hackers who distribute unauthorized copies of intellectual works. These techniques do nothing against agents, acting behind the veil of respectability, who invade our privacy and offend us by turning Cyberspace to their own interests and not ours.

We should take greater heed of the sanctioned harms of respectable insiders; we should question the systemic imbalances between the individual citizens of the online world, and the organizations that create it with little sense of the interests of the individuals. For the vast majority of Net users, it is the second group and not the first that is the significant danger; it is the second at least as much as the first, that affects our attitudes of trust online. Powerful security mechanisms may keep us safe from malicious outsiders at the cost of our online experience, but such mechanisms still leave us vulnerable to these agents. We can keep out the aberrant individuals, but we remain powerless against those parties that are poised to systematically exploit their positions. If we care about developing a climate of trust online—full-blown trust, not a thin substitute—we must address these conditions of imbalance between individuals and institutions. Evil hackers are not the only, nor are they the most important barriers to trust online. If we do not address the systemic problems, trust will erode and we will not easily recover from a sense of wholesale exploitation.

#### CONCLUSION: STRUGGLE FOR THE SOUL OF CYBERSPACE

I am not opposed to computer security. Underlying security mechanisms are diverse and can be shaped in an enormous variety of ways. The specific shape they take can have a significant impact on the shape of the online world and the experiences possible within it. Security technology does not, in general, necessarily lead to the state that appears at the end of the current trajectory, which would limit the complexity, richness, intensity, and variety of experience online without assuring us protection from a range of sanctioned predatory activities. Developed wisely, computer security may be able to produce a measure of safety with sufficient degrees of freedom to nourish trust. Cryptography is a good example: it can be used in service of transparent identification, but it may also be used to protect individual interests in privacy, freedom of association, and free speech.

Yet even the security mechanisms I have questioned here, namely, those that

enable surveillance, sustain Identifiability, and form selectively permeable fortresses—let us call this “high security”—are not in themselves objectionable. High security is good; it is even necessary for a great many settings: for airplane flights, military compounds, national secrets, nuclear power plants, banks, prisons, and more; all places where we welcome Richard Posner’s vaunted certainty.<sup>83</sup> If the arguments of this paper have succeeded, they will have convinced readers that the pursuit of trust must be decoupled from the pursuit of high security; trust will not ride in on the coattails of security. But these arguments do not in themselves provide an answer to the further question of what course of action we, the virtual agents, people, institutions who populate the online world, or they, influential parties involved in building and governing the technical infrastructures, ought to pursue. This we must recognize as a value about what we envision for the online world.

A highly secured cyberspace offers a good climate for activities like commerce and banking, and for established commercial, public and governmental institutions. The interests and modes of interactions that would not flourish are likely to include the creative, political, unusual, freewheeling, subversive, possibly profane, possibly risky modes and activities of individuals. For airplane flights, we may welcome security checks, but for these kinds of activities and interactions, for the virtual hustle and bustle that has come to resemble (and in some cases replace) much of our common experience, people avoid brightly-lit scrutiny. To express the tradeoff in terms offered by Luhmann, we may say that while both trust and security are mechanisms for reducing complexity and making life more manageable, trust enables people to act in a richly complex world, whereas security reduces the richness and complexity. Which one of these alternatives ultimately characterizes the online world—a virtual Singapore or, say, a virtual New York City<sup>84</sup>—should be a matter for full and deliberate consideration and should not follow merely as an accidental consequence of immediate technological imperatives and hasty policy choices.

In holding fast to the progressive social vision of Cyberspace, my choice would be an insistence on preserving the degrees of freedom that trust needs, while continuing to support a path of technical development in which security would ideally aim both to construct pockets of high security and to maintain minimal protections—safety nets to prevent catastrophic harms—in the realms outside of these pockets. If we set these as our goals, then we will have set the stage for trust. But we will *only* have set the stage. The work of nourishing trust and trustworthiness remains and calls for a familiar range of complex responses, including the promulgation of norms, moral and character education, and comfort for the hurt.

---

<sup>83</sup> See *supra* note 65 and accompanying text.

<sup>84</sup> I am grateful to Beth Kolko for suggesting this metaphor. When I presented a version of this paper in September 1999 at The Netherlands Royal Academy of Arts and Sciences a thoughtful audience argued that Amsterdam served better as a contrast.

2001]

*SECURING TRUST ONLINE*

131