Distributed Consolidation:
Identity, Reputation, and the Prospects for
Online Social Interaction

by

Alex Rosen
Class of 2008

A thesis submitted to the
faculty of Wesleyan University
in partial fulfillment of the requirements for the
Degree of Bachelor of Arts
with Departmental Honors from the College of Social Studies

Middletown, Connecticut        April, 2008
# Table of Contents

*Acknowledgements* ................................................................................................................. 2
*Introduction* .......................................................................................................................... 4

## I. Theory............................................................................................................................ 13
1. Trust............................................................................................................................... 14
2. Identity and Reputation................................................................................................. 19
3. Signals: From the Wild to the Web............................................................................ 30
4. Game Theoretic Model............................................................................................... 34
5. Conclusions.................................................................................................................. 50

## II. eBay........................................................................................................................... 52
1. eBay's Reputation System......................................................................................... 56
2. Problems...................................................................................................................... 67
3. Conclusions.................................................................................................................. 72

## III. World of Warcraft................................................................................................. 75
1. World of Warcraft's Reputation System................................................................. 82
2. Problems...................................................................................................................... 92
3. Conclusions.................................................................................................................. 97

## IV. Facebook.................................................................................................................. 101
1. Facebook's Reputation System............................................................................... 108
2. Problems...................................................................................................................... 115
3. Conclusions.................................................................................................................. 121

## V. Reputation Consolidation...................................................................................... 125
1. Why the Internet Needs a Consolidated Reputation System.............................. 125
2. Emerging Consolidation Standards and Their Problems..................................... 131
3. The Distributed Oracle............................................................................................ 147
4. Problems...................................................................................................................... 155
5. Conclusions.................................................................................................................. 159

*Conclusion* .................................................................................................................... 163

*Endnotes* .......................................................................................................................... 169

*Works Cited* ..................................................................................................................... 178
ACKNOWLEDGEMENTS

Like a child, it takes a village to raise a thesis, and this one is no exception. Now that it is all grown up, I'd like to take this space to thank a few of the parties to whom I am most indebted, though I will no doubt forget a few—to those, I apologize.

First and foremost, I express my deepest gratitude to my advisor Professor Richard Adelstein. His wisdom in all topics, thesis and otherwise, was invaluable over the past years. From the Sophomore CSS Economics Tutorial to talks in his office this spring, Professor Adelstein has always been willing to impart his knowledge and offer advice. He knew when to push, when to back off, and when to give praise. This thesis would not have been possible without him.

The Surdna Foundation provided generous support in the form of a Davenport Study Grant to start work on this thesis over the summer. Eddan Katz and Yale University supported my attendance at The Reputation Economies in Cyberspace Symposium at Yale Law School, a conference that proved to be particularly useful in formulating my ideas and exposing me to academics in the field.

Every professor and mentor I have had over the years deserves recognition. I am particularly indebted to Professor Gil Skillman, who helped me work through much of the game theory and was all too ready to poke holes in my developing argument. Thanks also to Professor Cherie Steele, who gave me my first opportunity to academically engage with the Internet in her CSS Junior Tutorial. Professor Cecilia Miller deserves my gratitude as well for working me to the ground sophomore year, always keeping her door open for advice, and generally being the first person that comes to mind when I think of my academic career at Wesleyan. I would also like to acknowledge Professors Rosenblat, Fay, and Paquette.

I am grateful to the College of Social Studies in general, whose students and faculty have made my Wesleyan experience what it was. Every class and late night discus-
sion has been formative in my intellectual evolution, and for these my gratitude is beyond boundaries.

Any of my friends who has ever endured my wild enthusiasm for some new website I had just found also deserves recognition, and probably my apologies. To Conny especially, who has been part of more of these episodes than anyone else, I give my sincerest thanks for more than I can express.

Finally, a big measure of appreciation to my parents, Jim and Carol Rosen, for reading the initial draft of this thesis and spending time on the phone talking through ideas with me. My father deserves credit for helping me choose this topic, even though he may not know it, by coming home everyday after work for as long as I can remember and talking about online identity management and Internet technology in general. I would probably not care half as much about this stuff if it weren't for him.

9 April 2008
Middletown, CT
INTRODUCTION

In a popular music video that circulated on the Internet recently, Jason Alexander (of George Castanza Seinfeld fame) plays a nerdy pizza delivery boy. The video opens with him being rebuffed by a customer, and as he drives off in his beat-up car, he turns on a live-video feed of Brad Paisley, the country music artist who is performing the song behind the video. The scene cuts to Paisley, who is performing a concert for thousands of screaming, adoring fans, most of them attractive females. He begins to sing, taking on the persona of the pizza-boy character, and describes his depressing life spent living at home with his parents and watching sci-fi. But, he continues:

There's a whole nother me you need to see, go check out My-Space.
Cause online, I'm out in Hollywood.
I'm six foot five and I look damn good.
I drive a Maserati, I'm a black belt in karate,
and I love a good glass of wine.
It turns girls on that I'm mysterious.
I tell them I don't want nothing serious...
I'm so much cooler online. So much cooler online.¹

The video ends with the pizza-boy realizing how silly his made-up online life really is, and he gets together with the girl he has had a crush on next door. But the underlying theme is clear: you can get away with anything online–even pretending to be a suave, rich country music star–though the result may be less than satisfying. Furthermore, the video suggests, those who we "meet" online may not be who they seem.

This paper is about how people evaluate others in online social interactions. Offline, it would be trivially easy to tell that the pizza boy is not a famous country
music star. Online, this can be difficult without the pudgy features, pizza delivery hat, and beat-up car as clues. Who knows, maybe the person with whom one is chatting really is Brad Paisley trying to connect with his fans. Then again, maybe that person is a child abductor posing as a kid's idol in order to secure a meeting. With a nightmare scenario like this so easy to imagine, it is a wonder that social interaction on the Internet is as popular as it is. How people build the trust necessary to connect with one another online—in diverse contexts such as to trade goods, ask questions, and play games—is the question that I examine here.

This question is also applicable to less silly and more common occurrences than a masquerading pizza-boy. After having received an email from a long-lost friend, how does one evaluate whether the person behind the email is really this friend? Why do eBay users give money to unseen sellers on the assumption that they will get what they paid for in the mail a few days later? How do players, known only to one another by their characters, decide to play together in massive online games that they take quite seriously? In short, how could anything possibly get done online, a medium that lends itself to anonymous, faceless communication and lacks a built-in way to identify users? The answer, I suggest, has to do with reputation. When users have some way to keep track of what the person hidden behind a pseudonym has done in the past and assess whether the information about themselves that they are broadcasting is true, it becomes easier to decide whether they are worthy of trust. This assessment does not have to reveal the person's real name or even particulars that correspond to his or her offline life, but rather the details that are necessary to build the trust required for the current interaction, such as whether the person has
shipped goods quickly in the past or seems to have the knowledge necessary to an-
swer one's questions.

Just as it does offline, online reputation serves a purpose. In every interaction, we make some judgment about reputation. Will this person keep a secret? Will that person pay me back? Will she get me there safely? These are some of the foundational practical questions that govern day-to-day life. They are not always noticed, however, because they have become so ingrained in daily routines. In all of these examples, though, the underlying question is about trust. Reputation is used to inform whether the information a person gives should be trusted, or whether they will do, or not do, something they say they will. Without this trust, people would find themselves paralyzed, afraid at every moment that someone will betray them. Reputation is necessary to support this trust because it forces individuals to consider the future when they are making decisions about how to act. A purely rational person operating in a world without reputation, as I show in the theory section below, will continually rip off whomever they are currently dealing with. For example, if a seller is not worried about his reputation, he will steal a buyer's money rather than relinquishing the paid for goods. If the seller is worried about how this will affect his future prospect for sales, however, he is more likely to cooperate and hand over the goods. I will formalize this model in the game theory section, and it will show how an effective reputation system supports trust by linking present to future payoffs.

The kinds of judgments that contribute to trust take place constantly in offline life, but with the relative luxury of being able to assess body signals, dress, and the opinions of others who may be close at hand. Reputation supported trust is similarly
necessary for online interactions, but deprived of these cues. Without face-to-face
cues, and often without any real knowledge of who the other interacting party is, trust
is devilishly difficult to maintain online and, like a rickety house in the middle of a
tornado, falls apart much more easily. Thus, reputation is left to uphold this fragile
trust and enable continued interaction. And whereas in offline life, reputation gener-
ally operates in mysterious ways, spontaneously arising whenever people interact, on-
line reputation systems are often consciously designed. Such control is both a bless-
ing and a curse. On the one hand, web designers are able to decide what type of trust
and purpose they must support and drop the correct reputation system into place. On
the other, as communist central planners found out the hard way, designing and con-
trolling complex systems like this is not easy and is prone to unexpected conse-
quencies. Luckily, websites have managed to develop ways to allow their users to
share identity and reputation information. The systems they have developed, as I will
show, have their benefits and drawbacks. I will also suggest that they could could up-
hold this task better, but as a whole they do a remarkable job of supporting trust in
their specific contexts among actors who may not know anything about each other.
This was not always the case; the early Internet lacked an effective way for users to
track one another's reputations.

The Internet was created as a way to connect the many existing networks of
the seventies and eighties. Once ARPAnet, originally funded by the Defense Depart-
ment, proved that connecting computers of different kinds over great distances was a
possibility, other networks, such as CSnet, BITNET, and Compuserve, began to
emerge.\(^2\) It was nearly impossible to contact users on other networks, whether to
send an email or share a computing resource. This was changed by the Internet, which, as Tim O'Reilly explains, is little more than a "network of networks" and consolidated the many networks that existed before its creation.\(^3\) The Internet, as such, is a set of protocols that enables computers to connect in order to share data.\(^4\) In doing so, it frees the identity and reputation information that users build up within these networks.

Early interactions on the Internet took place entirely in text, which meant that users had only this text with which to judge one another. Designers rarely built reputation systems into networks that allowed users to hold discussions, and users had to rely on different clues to judge others. On networks such as Usenet, which was organized into discussion groups structured around different topics, reputation was built entirely by posting history—which at first could not even be searched—and the name attached to postings.\(^5\) Those who spent a lot of time in certain newsgroups came to recognize certain pseudonyms and knew whether they had provided useful comments in the past or were new users trying to anger members for the fun of it.\(^6\) It was difficult, however, for new users to judge posters' reputations, and time and effort were usually the only remedy. Identity on the Usenet was mostly tracked by one's email address, which was attached to messages, and included the domain from which the users accessed Usenet.\(^7\) Certain domain names (such as those ending with .gov) proved a certain reputation, similar to the system that will be discussed in the Facebook chapter. In response to the lack of a reputation system, users developed their own mechanism based on diction and familiarity with certain conventions, a type of system that we will encounter again in the World of Warcraft case study. For exam-
ple, by making inside jokes in their signatures (a block of text automatically appended to the bottom of a posting), users could signal a reputation of understanding certain information. Similar mechanisms operated on other bulletin board systems such as the "Whole Earth 'Lectronic Link," but as a whole, these early networked groups generally lacked any organized way to keep track of reputation information. For the most part, however, they were small enough that this was not a problem, a fact that changed with the advent of the Web.

Even after the Internet consolidated the many networks that had sprung up, it was still difficult to find and link information on the Internet. In the early nineties, Tim Berners-Lee designed the protocols of the World Wide Web in order to make this task easier. Until the advent of the Web, most interactions on the Internet took place via a textual terminal environment, which meant learning extensive commands in computer codes such as Unix. This limited use of the Internet mostly to those studying computer science or very comfortable with advanced technical topics. However, the Web changed this, building a layer on top of the Internet that used it to link hypertext documents on computers all over the world in the easier to use environment of the graphical web browser that even Grandma can use. It thus consolidated the methods of posting and accessing information on computers connected to the Internet. With the World Wide Web, one can open up a web browser such as Firefox or Internet Explorer and type in an address, see an interesting hyperlink, click on this, and be taken to a new website with new linked information. By 1993, the Web had emerged as an accepted standard and consolidated the majority of Internet users onto the same system, thereby making the further connection of users (and data) possible.
consolidation created the need for reputation systems, such as eBay's, that would support anonymous users interacting with one another on websites, which could now be easily visited by anyone.

As the Web has matured, its designers have developed identity and reputation systems to support the trust needed for social interaction. A famous New Yorker cartoon published in the early nineties, showing a dog sitting at computer speaking to a dog on the floor, featured the caption "On the Internet, nobody knows you're a dog." As ways to broadcast and judge identity and reputation have spread, it has become harder and harder to pass as a dog in online social interactions. I believe, however, that this ability should be preserved while simultaneously making it easier to judge identity and reputation when it is needed. A dog, for example, should not be able to infiltrate a site for cat lovers. More seriously, political dissidents should be able to disconnect their professional lives from their political ones online, just as someone buying a camera on eBay should be able to require proof that the seller is not a Nigerian scam artist. I believe that fulfilling both of these requirements is possible by connecting the dispersed reputation systems that are now the norm online.

Just as the Internet consolidated networks and the Web consolidated access to information stored in different places, some way is needed to consolidate reputation systems so that they are better able to support trust. As it stands, one may have a reputation on Facebook, eBay, MySpace, Second Life, World of Warcraft, and SourceForge, but the information culled from each is not accessible to the others, even if they are referring to the same identity. They all exist within "walled gardens," their data unable to peek over the walls their designers have built. On Facebook, one's
connections (or "social graph") are not accessible on MySpace or LinkedIn. One's eBay feedback score, denoting that one is a trusted seller, does not follow one into Second Life, where one may also engage in the type of business that requires trust supported by an eBay-like feedback score. These measures of reputation are stored separately on islands of dispersed reputation floating in the sea of the Internet, making them, as I will show, more prone to fraud. Consolidating these systems, using the distributed oracle that I posit in the final chapter, has the potential to make reputation more effective by enlarging the group among whom it matters and providing users with more signals about one another to use in evaluation.

In short, my argument is that any website that relies on interaction among agents requires an effective reputation system to build trust that supports the site's purpose, whether this be supporting trade, engendering a social network, or providing entertainment. So far, sites have maintained their own siloed (closed-off) reputation mechanisms, but as the web evolves to become even more social, this walled-garden model of reputation will cease to work. A consolidated reputation system that further supports the web's strength for brokering cheap, context-specific ties must replace these many individual mechanisms. Without such a system, trust will continue to break down in the face of an onslaught of prisoners' dilemmas, leaving the Internet an ineffectual place for meaningful social activity.

The first chapter will outline the theory underlying these claims. Economics, philosophy, sociology, and biology have a lot to say about trust and reputation, and this chapter will take the reader on a tour of the highlights from these disciplines in order to impart the tools needed to understand the broader claims I make here. I also
formally model a game theoretical explanation of why reputation supports trust and cooperation by tying present to future payoffs. Much of the work in this literature review section does not directly relate to the Internet, but I do include Internet-specific research and attempt to raise signposts relating pre-Internet inquiry to this technology. Chapters Two through Four offer case studies of eBay (an online auction site), World of Warcraft (a massive, multiplayer online role-playing game), and Facebook (a social networking site). These services are emblematic of the three main types of interpersonal reputation systems in operation online today: explicit, performance-based, and implicit. They also all have marked weaknesses that illustrate the broad reasons why consolidation is necessary. Following the case studies, I examine emerging standards for consolidation and make claims as to what such a system, which I term the distributed oracle, must look like if it is to solve the problems of reputation and trust that mark the Internet today. Finally, a conclusion reviews the argument and makes suggestions for future research.
I. Theory

Trust, identity, reputation, and the relationship between them are concepts with great importance for both offline and online interactions. Because they are words often used without adequate definition, it is necessary to clarify them before being able to use them to discuss the case studies and the general importance of reputation supported trust on the Internet. The three are best thought of as of moving in a progression: one's identity feeds one's reputation, which supports the trust that is turn needed to support social interaction. How this works in practice is modeled by signaling and game theory.

In this section, I investigate some of the literature pertaining to these words as an infrastructure that will prove useful in this examination of online groups and the trust borne by reputation that permits them to function. I begin by offering a literature review and working definitions of identity, reputation, and trust. Signaling theory is then considered as a tool with which to analyze how people communicate identity and reputation, and apropos of online groups, how websites encourage this. Finally, I discuss game theory as a way to show why it is that reputation supported trust is so important to maintaining the quality of social interactions, especially in contexts where people do not repeatedly engage with one another. Together, this background will be important as a theoretical basis for the following chapters, where I examine different online groups and the ways in which their computer mediated reputation systems support, or do not support, trust. In light of the theory, these chapters
will illustrate that even though it may be surprising that any trust can be supported on
the Internet, a more consolidated reputation system is ultimately needed to ensure that
online social interaction can persist and thrive.

1. Trust

Trust is the basis of interaction. In a famous dialogue, Confucius outlines the
importance of trust to one of his disciples. Replying to a question about the pillars of
government, Confucius answers that food, defense, and the people's trust in their
leaders are the three most important factors in stability. When his disciple asks what,
as a leader, he should give up first, Confucius replies that weaponry should go first,
followed by food, because without the people's trust, "rulers will not stand."13 Thus,
trust is seen as the critical aspect in social cohesion, without which governments can-
not function; nor, as I hope to show, can groups as small as two.

Trust, at its core, demands a suspension of disbelief. It requires that agents as-
sum (sometimes based on past experiences) that other rational actors will do what
they expect them to do, even when the rationally selfish choice might be to do other-
wise. Furthermore, it requires that this assumption be made in the absence of any
iron-clad guarantees, a luxury which Onora O'Neill argues is rare anyway.14 As such,
trust stands in for legal contract, and when it is broken, so long as laws are not dis-
turbed, there is no legal recourse.15 Jonathan Cave notes that trust lowers costs in sit-
uations "where costs of contractual completeness are high or legal frameworks for
contract enforcement are unreliable."16 In this sense, trust requires a leap of faith, es-
especially in light of the mantra of investment advisors that "past performance does not guarantee future results."

As Robert Putnam writes, "trust entails a prediction about the behavior of an independent actor," suggesting the question of where the evidence for this prediction comes from.17 Trusting someone to do (or not do) something is a statement of confidence in their reliability. Nan Lin picks up on this leap of faith, defining trust as the "confidence or expectation that an alter will take ego's interests into account in exchanges. It represents faith that an event or action will or will not occur, and such faith is expected to be mutual in repeated exchanges."18 Extensive theory and empirical evidence has shown that continued group interaction and exchange leads to the strengthening of trust.19 In other words, the interactions and exchanges themselves, even divorced from the other benefits they might lead to, bring about a desirable outcome; namely, the strengthening of trust, assuming the interactions go satisfactorily for the person who comes to trust. Accordingly, even without a guarantee, trust enables people to place their confidence in certain outcomes, even when these outcomes are not explicitly defined or agreed upon. Whether it can be supported on a large scale in online interactions which lack repeat interaction is a question taken up by the rest of this chapter.

Given that one can never have an absolute guarantee that the person being trusted will justify this confidence, it stands to reason that people apply a threshold of

---

* Though whether trust leads to good behavior or repeat good interaction leads to trust is a chicken and egg problem that has interested thinkers such as Emile Durkheim and Niklas Luhmann (O'Hara, 2005, p. 444).
trust: those falling below this threshold will not be depended upon, while those above will command one's trust. This is the basis of Diego Gambetta's definition of trust, which states that trust represents a person's "subjective probability" that someone (or a group) will "perform a particular action," even when that person cannot monitor it. Gambetta notes that labeling another "trustworthy" implies that cooperation is possible with this person. Because trust is subjective, it depends on an agent's point of view, such that one person might trust Person A to save him from a burning building while another would not trust Person A with the most inconsequential secret.

Michael Bacharach and Gambetta also note that trust is subjective, because people are naturally disposed to different levels of trust. It is also "context specific," such that one may trust Person B to pass along a message but not take care of one's child. Similarly, some people are quick to hand out trust, while others are slow to offer it, even to close friends. In short, though trust cannot be quantitatively measured, people do seem to harbor internal guidelines about when and whom they trust. How they decide to place this trust in faceless actors online without traditional social cues is a question examined in the case studies.

Thinkers throughout the ages have praised the necessity and benefits of trust for groups of all sizes, from two people all the way up to society, often commenting on the instrumental, as well as intrinsic, benefit to groups filled with people who trust one another. Nan Lin summarizes Barbara Misztal's argument on the three functions of trust; according to Lin, "it promotes social stability...social cohesion...and collaborations. In other words, its motive is to maintain a group or community." This in-
instrumental view is similar to Durkheim's bonds of solidarity. Georg Simmel goes so far as to say that,

> Without the general trust that people have in each other, society itself would disintegrate, for very few relationships are based entirely upon what is known with certainty about another person, and very few relationships would endure if trust were not as strong as, or stronger than, rational proof or personal observation.²⁴

In other words, the ties that trust affords are necessary to hold groups of people together, and a general trust in people, even if small, is necessary for day-to-day functioning. Or, in Putnam's succinct explanation, "trustworthiness lubricates social life."²⁵

As the Simmel quote above makes clear, trust plays a vital role in holding groups together. So far, the majority of the discussion has been about trust between two individuals, and while this is useful as a starting point, trust within a group is more complicated. Trust flourishes in certain types of groups while it falls flat in others, a basic observation that is made clearer in the case studies. Given that groups can be large enough that all the members do not know one another, it is meaningless to speak of the trust that members have in one another in the same interpersonal terms used above. Members of these larger groups (or in the case of most interest here, online groups where members do not see one another, or even often know one another's real names) do not trust every other single person in an embodied way. Still, it is meaningful to talk about a group norm of trust that operates such that members can have a reasonable expectation of how others in the group will act in certain situations,
even if they do not personally know them. Reputation, as will be seen below, also supports members' abilities to make this judgment about one another.

For example, if a graduate of a large university meets a fellow alumnus in an unfamiliar setting, he is more likely to trust this person even though the two have never met, because they are able to find common ground by, for example, discussing where they lived on campus or whom they may know in common. Simply by the virtue of the two being part of the same group, these commonalities sow the seeds of trust. This kind of impersonal trust is, of course, weaker than the trust that is slowly built up by repeat interaction between two people. Returning to the college example, the alumnus might quickly realize that his fellow graduate is untrustworthy for other, more pertinent reasons. In larger networks of impersonal trust, as well, it is easy for "one rotten apple to ruin the whole barrel." If the fellow graduate ends up robbing the alumnus, he will be slower to place this perhaps unwarranted trust in other graduates in the future. Similarly, two people interacting on a website where they have had positive experiences in the past will be more likely to trust one another than if they interact on a site they have never visited. In sum, group-wide impersonal trust operates more fragilely than the trust between two close friends, but it is a property that is necessary for groups to function. This seeming contradiction is one of the fundamental problems that must be overcome in explaining how online groups whose members never see one another are able to function at all.

Trust is a critical part of any relationship or group activity, and it would be difficult to imagine the world working as it does without some measure of trust operating at all times. It is the expectation, whether rational or irrational, that someone
else is going to act in a certain way, often for one's benefit, without any contractual guarantees that this will happen. In this way, it fills in the gaps left by legal codes (which could not possibly be extensive enough to cover everything that trust does) and the market (because not every interaction is regulated by the price mechanism). Statements of trust can be as diverse as "I trust her to give me a fair share of the profits we make together" to "I trust that he is telling me the truth about who he is." People make these judgments every day using subconscious cues and a wealth of past experiences. While John Clippinger notes that the Internet holds the potential to "scale-up" trust among large groups of actors, it also lacks the information ascertained through face-to-face and repeat interactions, cues on which many classic definitions of trust rely. How to send and interpret these trust signals without the benefit of physical interaction is the intellectual problem this thesis examines, and its resolution that it only requires reputation, rather than the subconscious cues relied on in everyday life, is a surprising one. Identity and reputation—the factors to which I now turn—emerge as the vital component that supports trust among agents, even those previously unknown to one another.

2. Identity and Reputation

Identity and reputation are two sides of the same coin, connected by the way they inform and make statements about each other. Reputation is "meta-identity;" that is, a statement that uses identity information to describe attributes of a person such as "trustworthiness" and "honesty." As such, it becomes part of the person's identity, but because it broadcasts this meta-information rather than the factual data
that makes up a person's identity, the two are usefully treated separately. To make
this distinction more clear, I will examine identity, reputation, and the connection be-
tween the two.

For the purposes of this examination, a person's identity is composed of large-
ly factual information. Identity does not include the psychological components that
are important in how people understand themselves. I am not interested in how the
Internet mediates and contributes to the creation of users' identities, but more in how
these users prove certain aspects of their identities online, where face-to-face cues are
largely unavailable. In this way, identity as it is used here is more explicitly external
than the psychological sense in which the term is sometimes used. It is who a person
is, and what that person does. It includes things like one's name (or pseudonym), lo-
cation, age, gender, and job, as well as permissions such as whether one has the right
to use certain resources or post in certain places. One's identity can also include the
things that one has done in the past. In the context of a message board, this history
might include the past postings attached to one's username. Note that reputation com-
ments can be made about many of these: Is one good at his or her job? Does one re-
spect defined boundaries and only use what one is entitled to? Does one usually
make useful comments? In this sense, identity encompasses the *nouns* that can be at-
tached to a person, whereas reputation includes *adjectives* that describe these nouns.
Discerning these nouns is difficult enough in offline life, where we often rely on
physical cues to tell whether a person is lying, knowledge of what one is claiming to
be part of, or the assurances of others. Online, they become even harder to validate,
especially because the architecture of the Internet was designed in such a way as to be decentralized and make it difficult to track users.*

In computer science, identity information is most often used to authenticate users into systems. An identity is made up of claims about one's attributes, and based on these attributes, systems can make decisions about what a user has the right to do. A simple example is email. One chooses a username and password when signing up for a service such as Gmail. To Gmail, these are one's identity, though other identity attributes may be assigned, such as time-zone, in order to serve the correct information (in this case the time) to the user. The username and password, however, are the primary identifiers. Note that one's username identity does not have to correspond to one's real, offline identity. The system is just as happy with "kickin-babe989" as it is with "alexmrosen." Strong authentication, offered by companies such as VeriSign, promises to provide proof that users possess the attributes they claim (such as name). Another way to strongly authenticate users is by requiring them to pick up a username and password in person, but even then, these credentials may be stolen. Because these systems are costly, cumbersome, and often beyond what is needed to support online interactions, I do not examine them in any depth here.29

* These were very conscious choices, and though it would be possible to change the architecture to make it easier to track users by their real names (such as by putting an identity stamp on all their online activities), this would compromise the Internet's characteristic of enabling anonymous interaction when it is desired. Such a change, I believe, is neither attractive nor feasible, especially given the amount of protest it would face.
This notion of identity as authenticator is most useful in enterprise (business) contexts, where an automated system must keep track of a large number of users. Sometimes other humans can see these attributes (such as when I send an email from my Gmail account or sell an item on eBay under a certain username), but they almost always have to be vetted by some automated system, generally by a secret password. Much of the time, however, when users interact through systems such as email, eBay, World of Warcraft, or Facebook, they find it useful to signal other identity attributes besides those that can be authenticated. These might include age, time using the service, and location. The way that these pieces of identity are promulgated will be taken up more fully in the signaling section below, and the case studies will delineate some of the models that enable users to signal identity online, none of which have the power of my distributed oracle to share identity information among different contexts to support trust.

While identity is made up of the nouns that describe claims about a person's attributes, history, and permissions, reputation distills these into adjectives such as trustworthiness, skill at a certain act, consistency, and honesty that allow people to make and store judgments about others and their expected behavior. In other words, reputation is broadcasted meta-identity. Furthermore, while you know the most about all the layers of your identity (or identities), you can only read your reputation in the stories others make about you. As Bob Blakley, a principal analyst with The Burton Group, a consulting firm that focuses on technical identity issues, writes "your reputation is my story about you. You can't own this by definition; as soon as you own it, it's no longer my story about you; it instantly becomes an autobiography instead of a
reputation." As such, reputation is attached to identity, but just as offline, one can have multiple identities with multiple reputations. A party boy's family may think he is a perfect trustworthy angel from Connecticut, while his acquaintances may know him as a dishonest devil from NYC. Thus, reputation is a judgment about a person using identity information, but because this identity information may be inaccurate and multi-faceted, it is often a difficult judgment to make.

Abdul-Rahman and Hailes define reputation as a metric that informs "an expectation about an agent's behaviour based on information about or observations of its past behavior [such as gleaned from identity information]." Notice the similarity to trust, as reputation is a subjective belief, colored by one's information sources, and can be conceived of as a threshold agents use in deciding whether or not to trust other agents. One might live by the rule that one will not deal with any agent whose reputation falls below a certain level, though this is usually difficult to measure. Like trust, reputation is a belief about future behavior based on information drawn from the past. It describes general qualities one believes an agent to possess, such as trustworthiness or honesty, but may focus on other specifics appropriate to the situation, such as the likelihood that someone will pay back borrowed money, a reputation often summed up in credit scores. This highlights how a person may have multiple reputations corresponding to different portions of his or her life, just as one may trust someone else to do some things but not others. For example, one may have a great reputation for financial responsibility (as indicated by a credit score) but a horrible reputation for not being able to keep secrets (as indicated by the horror stories of
friends who have misplaced their trust). Thus, reputation is also context-specific and subjective.

Reputation, other than in quantifiably measured contexts such as credit scores, primarily depends on what Chrysanthos Dellarocas calls "word-of-mouth." For as long as humans have griped about a sale gone bad or a friend's broken trust, the most efficient avenue of spreading reputation information has been conversation. Whether complaining about the quality of a trader's goods in the Roman agora or grumbling around the water cooler about a new hire's inability to get the job done, conversations about others spreads reputation like wild fire. This information can quickly move beyond one's smaller social circle by being passed along from friends to friends-of-friends and so on. In offline life, this kind of spontaneous reputation system works remarkably well, because it is flexible and uses connections between people that already exist (their friendships and acquaintanceships) to spread reputation. Because a reputation is discerned from and attached to a corporeal person, it is often easier to judge. This is due to humans' adeptness at interpreting body language, a facility that is hindered when interacting online. Of course, as reputation travels through connections, it is prone to mutation through the "telephone game" effect. As each person, further and further removed from the identity information from which reputation is ascertained, adds their own interpretive lens, the reputational judgment is skewed.

This is why we, according to Robert Putnam, trust information that comes from close friends (or "strong ties" as he calls them, following Mark Granovetter) more than

* Companies can, of course, have a reputation as well, but I focus on human reputation here.
that which has travelled through many links. Spontaneously arising word-of-mouth systems take advantage of these natural connections for spreading reputation.

Avner Greif discusses one of these naturally arising reputation systems (methods of measuring and communicating reputation, even if crude and qualitative) in relation to medieval merchant guilds, where trading with foreign partners was supported by word-of-mouth reputations. The problem he examines relates to the fact that medieval merchants had to use agents in overseas ports to sell their goods. It was, for reasons more fully explicated in the game theory section, a rational choice for the agents to misreport sale prices so they could keep more profit for themselves. Thus, merchant and agent relations ran into a classic asymmetric information problem of the kind that George Akerlof examined in "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," and that is treated further below. Recognizing this problem, the Maghribi traders began to spread reputation information about the agents, such that untrustworthy agents found it difficult to get work. The merchants eventually formed a loose coalition, with the rules that only members of the coalition could work together and that those agents with bad reputations (i.e. those who had cheated a coalition merchant in the past) were not to be used in business dealings.

Reputation information was still traded by word of mouth among coalition members, but formalizing its use made it more effective. This success resulted from the increased broadcast power the coalition gave to reputation. In other words, the more people to whom one's reputation matters, the more powerful it is in governing one's actions. Greif finds that the merchants' system was quite effective:
This reputation mechanism explains the observed "trust" relations among the traders...the Maghiribi traders established a relationship between past conduct and future economic reward. As a result, agents resisted the short-term gains attainable through deception, since the reduction in future utility resulting from dishonest behavior outweighed the associated increase in present utility.\(^\text{39}\)

Because merchants and agents did not always meet face to face, reputation information was needed to signal trustworthiness in the absence of personally examinable identity clues. Thus, Greif offers an example of a group among which the need for reputation arose spontaneously but which quickly realized the need for a more refined system that allowed reputation to spread as far as usefully possible, in order to heighten its capacity for reward and punishment. The games (prisoners' dilemma and coordination game) that model this type of interaction will further make clear the importance that reputation plays in linking present and future utility.

As Greif shows, besides developing spontaneously, reputation systems can also be designed, a choice that the Internet and World Wide Web more easily enable than the word-of-mouth systems that dominate offline life. Because every bit of every website and program has to be programmed, designers must code in a way for agents to signal identity and broadcast reputation. Dellarocas posits that online reputation systems improve word-of-mouth networks by enabling them to operate on a larger scale at lower cost (in terms of time and money spent evaluating reputation) and offering "the ability of their designers to precisely control and monitor their operation," but encounter problems of volatility without face-to-face cues to help interpretation.\(^\text{40}\) Fred Pipier, Matthew Robshaw, and Scarlet Schwiderski-Grosche note that,
"In the traditional non-digital world we are constantly performing very sophisticated identification...[and reputation] decisions without realizing it. Our eyes and minds are remarkably adept at recognizing visual cues...yet, in cyberspace, we are robbed of this remarkable skill...." While online systems have the ability to broadcast one's reputation to more people than would be possible otherwise, the inflexibility and relative anonymity of the systems can be a fault and make them more susceptible to fraud. At the same time, because they may increase the number of people one can stay connected with beyond the anthropologist Robin Dunbar's number of 150 (the rough theoretical limit to the size of groups that humans' brains can keep track of), online reputation systems may increase the size of the "coalition" that Greif showed was so important to sustaining cooperative behavior among the Maghiribi traders. In other words, online reputation systems engender more efficient evaluation of others but are not without their problems.

As I mentioned, reputation is subjective because it is based on one's sources of information. The fact that it uses others' information and observations is what makes reputation so powerful. It allows evaluation of others without repeat interactions or even the expectation of future interaction. This is how one can trust someone one has never met before, based purely on a trusted friend's recommendation. Even if an objective or quantitative score is not available for most types of reputation, it still helps distill massive amounts of information that would be prohibitively difficult for any one person to gather. It takes advantage of "crowd-sourcing," relying on large groups of people to find and pass along information (such as when a group of friends gossip and trade stories about another person), and holds the property of being able to
be passed along and attached to a person without any central clearinghouse. In this capacity, reputation is analogous to price and particularly helps signal information about people in social situations where price does not come into play such as when deciding whether to trust a fellow player's recommendations for battle strategy in World of Warcraft. Online services, lacking avenues for traditional social cues, have attempted to centralize reputation in order to make it easier to measure and access. Whether there is a central authority or not, reputation makes it easier to decide whom to trust and whom to avoid, information without which social interaction would be prohibitively difficult. It thus simplifies interpersonal relationships and cooperative decisions in a world that, as the section on trust shows, often offers little in the way of guarantees or airtight legal contracts.

Reputation would mean little if people did not care about it, but because most do, it becomes an important motivator in social interaction. People take reputation seriously, because of the effect it can have on others' opinions of them and the doors it can open and close. This motivation suggests the second side of reputation. A person can have a reputation not only for being trustworthy but also for being erudite, funny, etc. Furthermore, having this reputation is desirable enough that in certain groups the free rider problem and other cooperative dilemmas are overcome. Because reputation is normally attached to a single identity—though this can be complicated when discussing an online reputation because it can be shared by multiple people or used by a single person as just one of many pseudonyms—previous acts may be difficult to escape and therefore agents have a motivation to "protect" their reputations. Similarly, when promises are kept or good deeds done, one's reputation rises in
stature, and it is this phenomenon that begins to explain the selfish motivations of altruism. As Judith Donath explains, "selfless goodwill alone does not sustain the thousands of discussions [online]; building reputation and establishing one's online identity provides a great deal of motivation."46 If helpful acts are seen as such by the group or community one belongs to, or if the receiver of the good deed spreads the word, one's reputation is bolstered.47

Both identity and reputation are critical to the functioning of any social interaction; without knowing something about with whom they are interacting, agents are poorly equipped to decide how to approach interactions. As stated previously, while identity encompasses nouns about people, reputation makes up the adjectives: whether they are good at something or trustworthy. In this sense, reputation is both a non-price signal and a motivator in and of itself. Being known as good at one's job can provide an ego-boost, but it can also lead to a promotion, while being labeled trustworthy can lead to others feeling more comfortable dealing with one, but also give one the self-satisfaction of being a "good" person.

Reputation and identity systems, whether spontaneous or designed, are needed to promulgate this information. In certain types of interactions (such as when engaging in an economic trade) it is useful to have a reputation system, while in others (such as on a dating website) it may be useful to have more in-depth identity information with which to make one's own interpretive judgments. Online reputation and identity systems help scale up reputation and identity broadcasting, potentially enabling trust to operate on a larger scale among agents—even pseudonymous agents—by increasing the size of Greif's "coalition." The information can still be faked, and it is
potentially easier to do so online where agents lack the face-to-face cues they are used to, leading to the conclusion that social interactions should be impossible online, though this is of course not the case. Because this danger can lead to misplaced trust, it is critical that the systems are designed correctly, and we will examine some of the current models, their faults, and potential remedies later in this thesis. Before investigating some examples, I turn to signaling, which has the potential to communicate identity and reputation information about oneself to others.

3. Signals: From the Wild to the Web

The original work on signaling came from evolutionary biologists, such as Amotz Zahavi and John Maynard Smith, who were seeking to explain why animals often waste resources that they should logically conserve. The answer, termed by Zahavi the "handicap principle," turned out to be that "wasting" a resource is a reliable way to demonstrate that an animal possesses the characteristic being signaled. According to Judith Donath, handicap signals are "inherently reliable, because producing the signal requires possessing the indicated quality." For example, when a gazelle stots, jumping up and down when faced with a predator rather than immediately running away, it reliably signals strength and energy, because the gazelle is able to waste these resources, confident that it will still be able to outrun the predator. An expensive car wastes money and signals riches and, as Veblen discussed, learning esoteric knowledge such as dead languages wastes time and signals a preponderance of leisure time, and therefore wealth. Conventional signals, on the other hand, are not as reliable as handicap signals, because they do not require possession of the attribute
being signaled. For example, wearing workout clothes does not prove that one is in good shape, and typing that one is female online does not prove this gender. These do not imply that the signaler has borne the costs that are to necessary to create a reliable signal. The preponderance of conventional signals, especially online, suggests that humans are gifted at faking signals, thus requiring a closer look at the systems that enable these signals.

Signaling is also an important concept in economics. Akerlof's "The Market for "Lemons:" Quality Uncertainty and the Market Mechanism" is the defining economic work on the problem of information asymmetry in economic transactions. His model showed that, in a market where goods are of uncertain quality, low quality goods (lemons, in used-car parlance) will drive out those of high quality, because there will not be a way for potential buyers to distinguish the two. Buyers, aware of this, will not purchase the goods for fear of getting a low-quality good, a cycle which eventually results in no trades at all. His model can usefully be extended to other types of human interaction where one party knows his or her own identity information, but the other party does not, creating an information asymmetry that results in an inefficiency such as the possibility of fraud.

Michael Spence, in response to this problem, developed his theory of "market signaling." He uses the example of the job market, where candidates signal their utility to potential employers with costly things such as an educational degree, which falls neatly in line with Zahavi's handicap principle. Signals such as these reduce the informational asymmetry between the applicant and hirer. Clippinger gives the example of a bank building an expensive office in order to signal that they are plan-
ning on staying in business for a long time and not running off with depositors' savings. Without reliable signals, participants are left to use conventional signals that can be faked, and therefore the informal asymmetry persists (the applicant, but not the employer, knows that he is a poor candidate; and the bank, but not the depositor, knows that it is planning to steal the money), which leads to suboptimal outcomes. Thus, signaling is needed to give the parties more information and build trust that they are getting a "high quality good," which in social interaction may be extended to mean a trustworthy person. We will see that reputation is one way to accomplish this.

How can reliable signals be maintained online to reduce this information asymmetry in the absence of face-to-face contact, when faking signals becomes simpler? It is, to restate the problem, easy for someone to write on an online dating site that he is sweet, caring and enjoys staying home and watching movies, when in fact he is aggressive and enjoys kidnapping his dates. Signals also lose reliability when other people begin to fake them. At the core, this resistance to fraud has to do with costs. If the benefits of faking a signal outweigh the costs, people will do so. Punishment is one way to increase the costs. Donath gives the example of putting a police siren on one's car to speed through traffic, an easily faked conventional signal, but one with a punishment that outweighs the costs for most people. The prevalence of conventional signals online leads to the need for reputational punishments for lying, as well as other ways to encourage reliable signals.

Donath and Danah Boyd discuss how the connections that people display can result in reliable online signals. They note that, "A public display of connections can be viewed as a signal of the reliability of one's identity claims," because links imply
that connections have vetted the information, whether it is on a personal website (to which others link from their pages) or a social networking service (one which others link directly to one another). Like offline friendships, connections are reliable signals of identity and reputation information, such as one's friends and social group affiliations. At a large enough scale, as when people have hundreds of friends on a service like Facebook, the system becomes prohibitively difficult to defraud, because one must convince many people to become one's online "friend," and usually--though not always--at least one person will catch on to the fakery and alert others or the person whose identity is being appropriated. Lying on a profile also becomes less likely, because one knows that connections will see the lies. So does inappropriate behavior, because others (whom one attacks or directs lewd speech at) know how to tell one's contacts, just as in the offline world. In other words, "by publicly displaying connections, one provides others with a means of getting in touch with one's circle of friends and acquaintances." Finally, a person with a connection to a trusted friend may be a signal of their own trustworthiness, especially if one counts the friend as a good judge of character. Revealing private information can be understood by the handicap principle as wasting the right to privacy in order to signal trustworthiness; thus public connections emerge as reliable signals of one's identity and reputation.

I submit that online identity and reputation systems emerge as a possible remedy to information asymmetry by enabling reliable signaling. Given an effective sys-

* It does happen. Recently, Bilawal Bhutto Zardari (the late Benazir Bhutto's son and successor as her party's chairman) was impersonated online, causing great detriment to his family and party (Nizza, 2008).
tem of management, reputation can be a reliable signal rather than its easily faked conventional cousin. If online identities take time and energy to develop, making exit costs high, the reputation is less likely to be discarded and more likely to be trusted; whereas if pseudonyms are cheap and have little history attached to them, trust will be less forthcoming. If an honest reputation results in benefits for its owner, and a dishonest one in punishments, identity signals are more likely to be honest. Furthermore, as discussed earlier, the farther reputation reaches and the more public it is, the more useful and reliable it tends to be. Thus, a consolidated system that amalgamates disparate identity signals (such as age, location, connections, and interaction history) into one reputation measure is preferable to one that only makes use of limited signals. Clippinger asks what could be deemed the primary questions relating to signaling, reputation, and trust: "How do you create the conditions for socially constructed and enforced honest signaling? How can reputation signals be credibly communicated and authenticated?" The case studies will address these questions by examining some of the current models that attempt to enable reliable signals, while the final chapter will propose a novel model of consolidating them to increase the amount of information available to interacting parties in order to support social interactions.

4. Game Theoretic Model

The tools of game theory offer a way to think about the concepts of identity, reputation, trust, and signals, especially in understanding why they are so important for sustaining social interaction. Game theory shows that when two people (players
in game theoretic terms) interact in a situation where each can choose to cooperate or not cooperate with the other, they often end up not cooperating, leading to an unsatisfying end result. This is especially true when the players know that they will only interact a finite number of times. Each will defect (that is, not cooperate) because they do not trust the other to cooperate. If they know that they will interact an unspecified number of times in the future, then each has a credible punishment (namely, not cooperating in the future) to encourage the other to cooperate in the present period. Of course, people do not always interact multiple times, especially online, so on the surface it would seem that all hope for cooperative behavior is lost. Luckily, reputation is a remedy to the situation, because one's reputation affects future prospects for cooperation, even if they are with other players, so that harm to one's reputation becomes a punishment. Reputation provides players with more information about one another on which to base decisions of whether or not to cooperate, potentially building trust among players as to how the other will play. In short, increasing the amount and quality of reputation information available to players will potentially lead to better outcomes, just as Greif illustrated in his medieval trade example.

Without any context, cooperation and defection are abstract words. For the remainder of this section, I will primarily use one example to ground the discussion. On the popular website eBay (a case study treated in depth below), buyers bid on sellers' merchandise, and the highest bidder wins the auction. After receiving payment, the seller sends out the item. An information asymmetry of the kind Akerlof described in "The Market for Lemons" results, because the buyer does not have as much information about the goods as the seller. Poor quality items may be represented as
high quality ones, creating the potential to drive out the truly high quality items (which result in less profit for the seller than the low quality items). I use a simplified model of eBay trade to transform it into a single-shot, normal form game. In this example, cooperation on the part of the seller means sending the item the buyer expects, while defection means either not shipping the item at all once payment has been received, or sending a lower quality item than is expected. Buyer cooperation means bidding on an item he or she wants and paying if the auction is won, while defection is not bidding in the first place or not paying for an item one has received. These concepts will be explained in further depth in the eBay section, but this surface understanding will suffice for now as an example.

Two types of games will make up the majority of discussion: the prisoners' dilemma and coordination game. The prisoners' dilemma is coldly obvious, yet maddening, in its principal lesson: rational actors will defect, leading to lower payoffs, even when they both know cooperation would lead to the better outcome (called Pareto optimality in economics) for both of them. This is because no matter what the other does in a prisoners' dilemma, each player will maximize his or her payoffs by defection. Looking at the game matrix will help:

---

* On eBay, the seller ships after receiving the buyer's payment, making buyer defection less of a problem. The simplified model, however, still serves our purposes by assigning real-world actions to abstract terms.

** For simplicity, player interactions are primarily modeled using a 2 x 2, or dyadic, game matrix. Real life is no doubt far more complicated, but distilling the game down to this least common denominator makes it far easier to discuss.
Figure 1: A basic prisoners' dilemma, with Nash equilibrium at strategy set (Defect, Defect).

The first number in each box represents the seller's payoff, while the second represents the buyer's. If the seller plays Cooperate (C), then the buyer will do best to play Defect (D) by not paying for the item and win eight, because he or she gets the item without paying for it. If the seller plays D, then the buyer will do best also playing D and end up with zero (because the buyer has not paid for an item he will not receive) rather than losing one. An outcome of (D, D) represents a "no trade," where the buyer does not even bother bidding, expecting that the seller won't ship anyway. In the prisoners' dilemma, (D, D) is referred to as the Nash equilibrium; in other words, the outcome that rational players will end up at by playing their self-interested strategies.

* Note that rational self-interested players' utilities may take into account feeling bad for the other player and caring about how the current interaction affects future ones. In the prisoners' dilemma, these are largely left out. An evolutionary, rather than a rational choice, model would also yield dividends for modeling player interaction, but is beyond the scope of this paper. In this model, identity might correspond to "type," which reputation signals. For excellent summaries of evolutionary play, see Axelrod, 1984; Hofbauer and Sigmund, 1998; Smith, 1982.
Peter Kollock puts it succinctly when he describes social dilemmas, of which the prisoners' dilemma is the most famous, as "situations in which individual rationality leads to collective irrationality." These abound in offline life, where their solutions have been studied extensively, and also creep into online activity.

Another famous prisoners' dilemma is the "tragedy of the commons" that Garret Hardin famously posited. The example he uses, that of a herdsman using a common pasture, gets at the core of the problem elegantly. Every herdsman knows that adding another one of his sheep to graze freely on the commons will increase his utility by giving more of his sheep free food, but if every herdsman does so, the commons become overpopulated and destroyed, leaving no resources left for anyone.

One of the other most common forms of this dilemma is the free rider problem. In this dilemma, each rational actor, perceiving that some goal will be met even if he or she does not help to meet it, ends up "free-riding," leading to the failure to reach the goal, such as when someone asks questions in an online discussion forum but does not answer others' queries. If everyone does this, the forum loses its effectiveness. In sum, everyone knows what would be best, but they still end up with the suboptimal outcome because of their narrow self-interest, an outcome that will be repeatedly encountered in the case studies.

The other game of interest is the coordination game. Social dilemmas modeled by the coordination game are less depressing, because they have two possible Nash equilibria: one where both players cooperate and the other in which they both
defect. The same example of eBay can be used here, but with the addition of reputation, such that defection has a reputational cost attached to it.*

* eBay does in fact include a feedback system (that we will encounter in the case study) with which players can rate each other on the transaction in such a way that other, future players are able to see it.

---

**Figure 2:** A coordination game, with Nash equilibria (Defect, Defect) and (Cooperate, Cooperate).

If the buyer thinks the seller will cooperate (and vice-versa), it is in his best interest to also cooperate and win five rather than the four he would win if he defected. But if the buyer thinks the seller will defect, he would do best to defect as well and win two rather than the one he would get by cooperating. In the eBay example, when only one player defects, the cooperating player is likely to give him poor feedback, resulting in a worse reputation and lower utility. A mutual defection may result in a no-trade outcome, but at least neither player risks losing reputation points. For this
coordination game to work, the assumption must be made that players also value their reputation (in addition to the payoff from the trade). A poor rating is a credible punishment because, we will assume, it hurts a player enough to make him want to exact reputational revenge when he cooperates and his opponent defects. In addition, valuing one's reputation implies valuing the future, or in other words, that the players have a "time-discount factor" of less than one.\textsuperscript{67} That is, their utility payoffs take into account reputation effects that they know will matter later on, implying that they plan to continue to have interactions with others who would use their reputation as a basis for interaction. Because of this effect, the Pareto optimal decision is not ensured in a coordination game, but it is at least attainable, unlike in the prisoners' dilemma.

These games suggest the need for trust. As discussed earlier, many interactions take place without the guarantee of a legal contract and therefore require players to rely on trust. Dellarocas explicates this point, writing that "At the heart of any bilateral exchange there is a temptation [for a player]...to defect from the agreed upon terms in ways that result in individual gains for it (and losses for the other party)...[u]nsecured bilateral exchanges thus have the structure of a prisoners' dilemma."\textsuperscript{68} This fear can be assuaged and the dilemma resolved with trust, which is often built up as a result of repeated interaction, resulting in a player's confidence that the other player will uphold his or her end of the bargain, which transforms the series of prisoners' dilemmas into a coordination game. Unfortunately, and especially online, players often do not meet more than once; for example, a seller on eBay may sell to thousands of customers, but to each one only once, making it difficult to develop trust in the more traditional ways that flourish when traders repeatedly meet face-to-face.
In these situations, reputation acts as a proxy for interaction history and the trust this produces by allowing players to make informed decisions regarding the likelihood that the other party will cooperate even when they have not met before. With reliable reputation reporting, it is necessary that players will keep playing the game, but unnecessary, as Nolan Miller, Paul Resnick, and Richard Zeckhauser write, that the "same partners" will interact multiple times. If players suspect that their opponents are not going to keep playing the game (for example, continuing to trade on eBay) where reputation is important, then a negative rating is not a credible punishment and the coordination game morphs back into a prisoners' dilemma. But, if players will continue to play, reputation, in Robert Axelrod's words, "enlarges the shadow of the future," so that retaliation against defection is a credible punishment to players who have a sufficiently large present value of future payoffs. In these situations, it is clear why it is in players' best interests to build up a good reputation and interact only with others who have a good reputation (as in Greif's coalition), thus reducing the transaction costs of continually evaluating new players in one-shot games where they risk the other party's defection. Players who trust others experience a lower level of risk aversion—the fear that the other player will defect, which leads to a play of defection to maximize payoffs against a defector—and therefore feel more comfortable playing cooperate when they think the other will do so.

Reputation thus acts as a form of Spence's signals that informs the other player of whether one is likely to cooperate or defect—a shirker of high quality goods or a fly-by-night fraud—and thus reduces information asymmetry through pre-play communication. Like a reliable signal, reputation can be costly to build up and main-
tain. Of course, defectors do not want to broadcast their true reputation, as the continuing history of online fraud has shown, and they prefer to signal trustworthiness, defect, and then repeat the cycle under a new, assumed name and reputation. Nevertheless, with a reliable reputation system in place that punishes liars, and that is used by players who value the future, the prospect of the other player reporting that one has played cooperatively serves as a payment. This encourages and rewards individual cooperation. As Gary Bolton, Elena Katok, and Axel Ockenfels write, "a reputation mechanism...makes trust and trustworthiness lucrative behavior." If one's reputation falls below some threshold, one might not be able to participate in future games and be banished from the website or group. Reputation is thus a signal that broadcasts past play (identity information) and puts players on more equal informational ground.

Reputation connects individual prisoners' dilemmas into a larger game, as shown below, in which cooperation is possible by linking the payoffs of the present game to future games' payoffs. Praising the prospects of this connection, Dellarocas writes that "if the promised future gains of reputation effects are high enough, ordinary sellers are induced to overcome their short-term temptation to cheat and to try to acquire a reputation for honesty by repeatedly delivering high quality." A similar notion can be applied to non-monetary games (such as when deciding whether or not to connect to a new person on an online social network) but discussing trade examples (in this case eBay) makes the idea clearer. Thus, players must take into account

* This is the problem of "cheap pseudonyms" that Friedman and Resnick, 2001 and Feldman and Chuang, 2005 have identified.
long-term effects on their reputation, and not just the short-term temptation to cheat. Of course, the signal that reputation will be reported by other players has to be reliable for this to work. If players think that they can get away with defection in secret or swap in for a new, low-cost pseudonym, then the signal will not be trusted. If it is, however, then reputation helps overcome the prisoners' dilemma to reach Pareto optimality, the best outcome for both players.

When adequate signals feed reputation, a series of repeated prisoners' dilemmas with the credible punishment of negatives reports may become a "coordination super-game." These games must have no known end, or else defection will occur just as in a single-shot prisoners' dilemma. Kaushik Basu explains this well in relation to a series of 100 prisoners' dilemmas:

Consider the 100th period. That being the last, each player has reason to maximize his or her earning in it. So they will play as if it were a single-shot prisoners' Dilemma. So the outcome will be (D, D). Now consider the 99th game. Since the outcome in the 100th game is a foregone conclusion, each player may as well play D in this one too. The argument unfolds inexorably backwards and the outcome is (D, D) in every game.

But if players do not know when they will stop playing, the repeated series of prisoners' dilemma may morph into the coordination game discussed above. George J. Mailath and Larry Samuelson note that "repeated interactions give rise to incentives that differ fundamentally from those of isolated interactions," namely that players will want to protect their reputations by not defecting. Again, players do not have to

* Note that cost includes the costs associated with a new pseudonym, such as the reduced trust or lower prices available to the seller in an online auction.
meet multiple times, because they know the reputation they accrue from interaction will affect future games with others. When reputation effects are considered in the payoffs, the coordination game that results means that cooperation becomes a Nash equilibrium.

In this coordination game, as has been shown, if trust is high enough to make an accurate prediction of the other player's strategy, then it is far simpler to decide on one's play. If one expects the other to cooperate, then cooperation is one's best option, while if one expects the other to defect, then defection is the best decision. It is, again, likely that players with poor reputations will be labeled defectors and find it difficult to continue playing the game and collecting payoffs—in the current example, the gains from a successful eBay trade. Because reputation links the individual games' payoffs to one another (by determining whether others are willing to play with one in future games), it results in the possibility of cooperation in each of these sub-games.

One problem, alluded to above, is that risk aversion may still take over, and the fear of the other player's defection may lead one to defect, even with the understanding that cooperation may be possible. Reputation-fed trust is also needed to overcome this situation. Russell Cooper discusses how pre-play communicative signals can overcome players’ risk aversion by giving assurances (even if they are not guaranteed) that the other player will play cooperatively. As empirical proof, he cites a coordination game experiment where pre-play communication led to 90% of the outcomes being at the Pareto-optimal level. Ultimately, according to Cooper, the equilibrium outcome will depend largely on the level of the players’ risk aversion.
Given a generalized game, such as the one below, Cooper solves for a variable $s^*$ that equals $(4-x)/4$ and represents the value at which a player is indifferent to playing Cooperate or Defect.

![Cooper's Game](image)

**Cooper's Game**

According to Cooper, “if a player attaches probability less than or equal to $s^*$ that his opponent will chose action 1 [Cooperate], then that player ought to select action 2 [Defect]” and vice-versa. Thus, how much risk is present, in the form of the relative values of payoffs in strategy sets (Cooperate, Cooperate) and (Cooperate, Defect), partly determines the outcome, but so does the value one assigns to the likelihood of the other player cooperating. With signals, this value can be increased. Notice the similarity in $s^*$ to Gambetta's definition of trust discussed earlier, which suggested that trust is the probability of another person doing a certain action. Trust that the other player will play cooperatively will thus raise $s^*$ and lower risk aversion, making cooperative play one's rational choice.
Reputation can support cooperation, but only if actors can easily identify one another. The more accurate this identification is, and the more easily that players can recognize one another's reputations, the more robust cooperation will be.\textsuperscript{82} Even if there is an effective reputation system, past behavior may not accurately predict future behavior, especially if defectors try to disguise themselves as cooperators. Luis Cabral notes that dishonest agents have the incentive to cooperate in early games in order to defect in later rounds to make a profit.\textsuperscript{83} Per Akerlof's theory, fraudsters have more information than their targets about when they are planning to stop playing the repeated games (and therefore stop caring about their reputations). For reputation to be an effective identifier to use in making strategy decisions, players also need to be sure that it has the power of credible punishment over their opponents.

As Greif discussed, enlarging the size of the "coalition" where reputation has an effect is one way to approach the problem of reliability. Of course, in Greif's example of medieval merchants, reputation only mattered in one context, that of trading. But because this context was so important, the "grim trigger" strategy of refusing to play with those who had defected in the past made an injured reputation a credible punishment. One way to increase the coalition's size is to make reputation matter in more than one context. For example, rather than just having one's eBay rating affect activity on eBay, it may affect activity in other contexts where trading is involved, even within a game like World of Warcraft where players buy and sell virtual goods. Linking online reputation to a social networking site such as Facebook, where one makes public one's friendship connections, may act as a reliable signal that one plans to cooperate, because the coalition now includes not only strangers but offline friends.
These are only a few examples of the kind of consolidated reputation possibilities that I will discuss later, but note how they expand the coalition's size. By providing players with more diverse signals about one another, a reputation system that links contexts is therefore better equipped to support trust, reduce risk aversion, and enable cooperative behavior.

Increasing the size of the coalition by linking reputation contexts has three necessary conditions to keep it reliable. First, players must be sure that their opponents get value from reputation in each context that is linked; otherwise a negative reputation will not be a punishment. Second, as discussed in the previous paragraph, there must be some guarantee that others will continue to play the game with the identity to which the reputation ratings are attached; otherwise, only they will know that they are in their last period, in which defection is the logical option. Guaranteeing this condition is difficult, and sometimes impossible, but for certain contexts that link to offline life (such as online social networks), it may be feasible. Third, opponents must not have an easy way to counter the reputation rating if it is negative; otherwise they could argue that it was a false rating and not lose the prospect of value from future play. These requirements are sometimes hard to assure, but players are able to take into account their estimates of the probability they think the conditions are satisfied. As the number of contexts grows, this probability will edge closer and closer to one as it becomes harder to defraud each and every context.\textsuperscript{84}

Reputation has, ironically, a commons problem of its own, and is thus plagued by the very trouble it is called on to solve. Gary Bolton, Elena Katok, and Axel Ockenfels conducted a study that compares the partners market (in which two individuals
continually trade with each other) with a reputation-supported market (in which many traders interact with one another using reputation as a guide). Finding that the reputation market performs well, though not as well as a partners market where two people continually interact, they find a distinct public goods problem within reputation that may be holding it back. Basically, supporting a general norm of trustworthiness—by cooperating, building up a cooperative reputation, and then continuing to cooperate—is a benefit for everyone in the market, because it reduces fear. They explain that the trust effects of reputation are only "fully internalized" when partners continually interact, but not when players contribute to a public reputation on which others rely.

Allowing reputation to be accessed across websites creates the potential for an even larger free rider problem: why should a website (or person) contribute to the score if it can just rely on other sites to do so? On the other hand, consolidated reputation may be so much more reliable than when it is siloed in different contexts that websites will want to contribute to it in order to reap its benefits. When the structure of the mechanism is designed correctly, the payoffs encourage websites to participate in the measure, as will be discussed further in the final chapter. Thus, reputation has a prisoner’s dilemma of its own, but one that can be overcome with the right kind of consolidation.

Game theory is a useful tool to structure how to think about interactions between people. Of course, it is simplified: no bounded set of strategies and payoffs can do justice to the complexity that underlies social interaction. Even so, its lessons are powerful as we move to discussing case studies of real reputation systems at work in online groups. The news constantly features stories about online fraud and stolen
identity, but when viewed through a lens of rational action, it is surprising that so much cooperation persists. One could also say this about offline interaction—why don't people steal when they know there will not be negative consequences?—but it is unexpected that reputation and trust can emerge without face-to-face contact and the threat of the immediate, in-your-face shame that marks offline interaction.

Luckily, game theory also provides the tools to explain this surprise. Repeated interaction leads to players valuing future payoffs, and when punishment in the form of lesser payoffs is guaranteed by reputation, short-term selfishness—the root of the prisoners' dilemma—is replaced by future-valuing forethought. The tenuous nature of this cooperation, however, should be obvious by now. Fraudsters enter the game, people change their strategies and begin anew with fresh pseudonyms, and players make mistakes that are interpreted as conscious defection. Reputational punishment that spans multiple contexts must toe a fine line between being too lenient, in which case defectors will multiply, and being too stringent, in which case players will stay out of the game for fear of unintended actions leading to harsh punishment. Consolidation thus has a difficult problem to overcome, but in its power to increase the amount of identity and reputation information available to players, it has the potential to engender cooperation far better than when reputation is only visible among a small group. The case studies engender cooperation far more than would be the case without a reputation system, but they will show that no one model has emerged as satisfactory for supporting cooperative trust that is immune to defectors, except for, perhaps, the promise of a consolidated system.
5. Conclusions

By being able to signal information about others' identities (most notably their histories of play), reputation is critical to building the trust required for social interaction. If people are unable to signal this information reliably, informational asymmetries will arise and social interaction will collapse into the pitfalls of single-shot prisoners' dilemmas. The way that identity and reputation is broadcast online is particularly critical for the prospects of online interactions, where participants lack the face-to-face cues that are so important in offline life. If the costs of evaluating others are too high, short-term social interaction will be impossible, because the parties involved will not be able to build the basic trust required to interact. One of the Internet's greatest powers is its ability to anonymously connect users across the world who may not know anything about one another. But this is also its greatest fault. Without a foundation of trust on which to interact, Internet users will find it difficult to answer one another's questions, trade goods, evaluate whether others are who they say they are, and collaborate on projects as small as playing a game together and as large as writing the Wikipedia encyclopedia. Reputation operates in varying ways within different types of groups, but all of them require some form of system with which members can evaluate each other and discern the identity information needed to support trusting interactions.

Based on the theory presented in this section, it would be expected that increasing the number of reliable signals and the size of the coalition within which reputations matters will reduce defection in social interactions by reducing informational
unevenness. In other words, being able to identify others is critical to building the trust needed to interact, but it is unclear whether it is possible to prove that a player's signals are truthful and relevant to all contexts. These theoretical propositions will be examined in the rest of the paper to show whether it is practical to enable this reputational identification online, and if so, what the most efficient way is to uphold it. In the case studies, I examine three reputation systems (eBay, World of Warcraft, and Facebook) that have emerged, their strengths, and ultimately why each of them is unsatisfactory. The final chapter suggests how a consolidated reputation mechanism is required to build the trust needed to transform repeated prisoners' dilemmas into a coordination super-game.
II. eBay

The background section showed that reputation supported trust is critical within any group in which members do not repeatedly interact with the same people. If trust is the ability to predict, to a certain degree, the likely future actions of others, then an absence of trust will wreak havoc upon social interactions. Fear takes over, leaving people perennially suspicious of one another and unable to enter into short- or long-term agreements or undertake shared projects that require mutual assurances that each will do his or her part. In short, without the signals to help build trust, lasting ties and cooperation are impossible.

These ideas hold especially true for online interactions. On any website that is centered around the contributions and interactions of its users, reputation plays a role in sustaining the trust that supports the site's purpose. Note that some websites attempt to define themselves by the notable absence of identifying information. Invisi-blog, an anonymous blog hosting service, boasts that, "You don't ever have to reveal your identity - not even to us. You don't have to trust us, because we'll never know who you are." Such a claim might be true, but blog authors still maintain a reputation based on their history of posting, even if this reputation is not linked to their offline identity. Other websites pride themselves on how accurate their members' information is. LinkedIn, a business-themed social network aimed at professionals, depends for its success on users' belief that those they are connecting with are who they claim they are and present real opportunity in the form of increased social cap-
ital. That is, on LinkedIn, reputation is linked to offline identity. The case studies presented here fall at different points along this spectrum, between anonymity and full disclosure, though all of the sites do technically have some way to figure out who a user is, even if one can get away with using a false identity.* On eBay, one is only likely to know the true identity of another user if one completes a trade with that user, because pseudonyms are used until it is time to reveal a shipping address. In World of Warcraft (WoW), users go by made-up pseudonyms, though they must pay monthly by credit card, implying that Blizzard, the company that runs the game, has their real identity on file. On Facebook, users are required to include their real name on their profiles—though of course this does not always happen—and often choose to include a picture and other identifying information.** Reputation may be linked to offline life more explicitly on a service like Facebook, but is no more important than it is on eBay and WoW.

In this section, I will examine the reputation systems in use on eBay, WoW, and Facebook, noting how the different systems support each site's differing purposes. In addition, for now I will treat each service as a "walled-garden" in which reputation does not spill over between sites. The reality is more complex, however, and the Internet is likely beginning to consolidate in such a way such that this will not always be the case, a topic that will be treated in Chapter Five. Together, the case stud-

* The companies are able to track a user's IP address to locate him or her, unless the user is advanced enough to circumvent this, an ability beyond the abilities of the average Internet user. Tracking IP addresses is difficult, and usually requires a subpoena to the ISP a user is on, but it is possible.

** Briefly, the service is significantly less useful, in the sense of being able to build and maintain a network, if false identity is used. This idea will be expanded in Chapter Four.
ies illustrate the importance of a reputation system to support signaling in social interactions, but they also show the problems that arise when this system is limited only to the site at hand, thus supporting the theoretical propositions made in the previous chapter.

****

eBay was founded in 1995 as an online forum where people could connect in order to buy and sell goods to one another. At its start, eBay, like most other early sites on the Internet, was small and intimate, and the first item sold is reported to have been a broken laser pointer. Available items numbered only in the thousands. According to David Baron, "The Posse," a group of six members, "actively monitored activity on the eBay site. Traders who violated implicit standards could quickly have their reputation damaged by The Posse...[via] the communication they exchanged through email and on bulletin boards." When the number of listings began to explode in number, however, eBay quickly realized that a more robust system was needed, leading its founders to establish the Feedback Forum, a reputation system that enabled buyers and sellers to rate one another on transactions and view others' ratings before electing to bid or enter into a trade. Thus began what is arguably the most well-known reputation system on the Internet.

Traders on eBay face an information asymmetry problem. Only the player planning to defect has this knowledge, so the other player enters the relationship at a disadvantage. In the way that Akerlof described, dishonest sellers will drive out hon-
est sellers because they get a larger payoff from a single-shot interaction. This is made worse by the fact that most of the transactions on the site happen between buyers and sellers who will only trade once and therefore lack the power to threaten the other with future defection. The theory predicts that without credible signals, a no-trade equilibrium would result. Why this is the case should be clear based on the model in the previous chapter that used eBay as the primary example. The theory also predicts, however, that an effective reputation system should mitigate this asymmetry problem and support interaction, in this case trading.

Early on, eBay put in place systems to support the trust for which the site is known. According to eBay, "Feedback [reputation] fosters trust between people by acting as both an incentive to do the right thing and as a mark of distinction for those who conduct transactions with respect, honesty, and fairness." A reputation system has helped develop trust and convince users that eBay is a safe forum within which to conduct real-money business. In theoretical terms, feedback acts as a signal to future traders, turning the series of iterated prisoners' dilemma games into a coordination super-game where cooperation is an effective strategy. It thus stands in for repeat interactions and addresses Akerlof's information asymmetry. This section examines whether cooperation has actually emerged.

Though eBay does not release specific statistics, it is clear that fraud has increased commensurate with the growth of the website. As eBay became too large for members to police, reputation has been left alone to support the trust needed to engage in trades, and the reputation system has been powerfully tested as to how well it supports eBay's purpose to make profit. Thus, reputation is left as the sole means to
support the trust needed to engage in trades. It should be surprising that this system has worked at all, given that reputation is most often passed on by word-of-mouth rather than a disembodied score. But, in fact, reputation has supported trust remarkably well in an environment where members have only the feedback score as a signal with which to judge one another. I will show, however, that the relatively small size of the eBay "coalition," following from Greif, has made it harder to support reliable signals.

1. eBay's Reputation System

eBay's reputation system aims to enable the type of trust that is found in traditional marketplaces where sellers and buyers interact multiple times and in face-to-face contexts. In fact, the central problem that prompted eBay to develop its feedback system was the fact that there are very few repeat interactions on the site. Paul Resnick and Richard Zeckhauser found that 89.0% of buyer-seller pairs traded only once during their study, a number likely to have grown since their 2002 study.\textsuperscript{91} Thus, trading on eBay does not enjoy the benefits of Bolton, Katok, and Ockenfels' partners market, discussed earlier, which is marked by repeat interaction.\textsuperscript{92} Still, eBay's reputation system mimics and spreads the type of information gleaned from repeat interaction by representing all past interactions (which constitute identity information) within an easily digestible index, the feedback score. In doing so, it supports a multilateral system of punishment, because one knows that if a deal goes sour, even though one might not have the chance to elicit payback in the future, a negative feedback score will signal others to carry out the punishment in the form of lessened prob-
ability of sale or reduced prices, as discussed below. Even though it would seem that the interactions are single-shot and therefore prone to the uncooperative pitfalls of prisoners' dilemmas, this creates the coordination super-game that links present to future payoffs. Because the system is not perfect, these pitfalls do trip up users, but the reputation system largely mitigates them by enabling traders to signal planned cooperation in order to build the trust needed to lower others' risk aversion.

Only after conducting a trade on eBay are the buyer and seller able to rate each other. Rating is done on a scale of positive, neutral (which is often viewed as negative, though not in the score), and negative, whereby a positive rating counts as +1, a negative as -1, and a neutral as 0. Per a recent change intended to reduce collusion that would render feedback signals unreliable, repeat ratings from a user only count once in the receiving member's feedback score. Ratings are reflected in a feedback score, which is displayed next to all members' screen-names, and different colored stars denote feedback levels. In the screenshot below, the star (which is turquoise when reproduced in color) represents feedback totaling between 100 to 499 points. This information is displayed next to the seller's name on each auction page, along with the percentage of positive feedback, how long the seller has been an eBay member, and the return policy offered, if any—all signals of trustworthiness that help buyers decide whether the seller is likely to cooperate or not.

* Other conventional signals do exist, such as the professionalism of the seller's auction page and tone in emails, but these are beyond the scope of this paper. Other ways of signaling reputation include guarantee policies and accepting credit cards (which reveals that the seller has been accepted as legitimate by the credit card companies and would be easier to track down in the case of fraud), but it is clear that the feedback score is the most used, and therefore most important, measure.
Meet the seller

Seller: ipodrepairpros (100 🌟)
Feedback: 99.0% Positive
Member: since Mar-15-00 in United States
- See detailed feedback
- Ask seller a question
- Add to Favorite Sellers
- View seller’s other items

Buy safely

1. Check the seller’s reputation
   Score: 100 | 99.0% Positive
   See detailed feedback

2. Check how you’re protected
   This item is not covered by buyer protection on eBay

   Returns: Seller accepts returns.
   7 Days Money Back

Figure 4: Some of the information displayed about the seller on an auction page.

Clicking on the member’s name or feedback score brings the user to an extended feedback profile, shown in Figure 5. This profile displays feedback totals for the last one-month, six-month, and twelve-month periods, so the user can get a better idea of when feedback was given and how trustworthy the seller has been recently, making signals of past cooperation more specific. There is also a section of Detailed Seller Ratings (DSRs), which allow buyers to rank sellers based on item description accuracy, quality of communication, shipping promptness, and accuracy of shipping
and handling charges. DSRs were introduced after the original feedback system as a way to let buyers offer more specific feedback on common areas of concern. Finally, users are able to view others' short messages, left in conjunction with their score and often along the lines of "Item as described. Quick shipping. Pleasure to deal with. A+++." Taken together, this information is supposed to provide an accurate picture of a member's history on the site and act as a signal that the member will cooperate in the future and is therefore worth trusting.

Figure 5: Part of a user's extended feedback profile.

A seller's feedback score is considered much more important than a buyer's score. Sellers can be relatively confident about dealing with buyers, because they can always wait for payment before releasing an item, complicating the simple trading game I posited in the game theory section above, though negative feedback may sig-
nal a buyer who backs out of won auctions.” Buyers must make more of a leap of faith, trusting that the item will come as described, or at all. eBay transactions are thus best modeled as an extensive-form game, as opposed to the normal-form games dealt with above, that captures the decision tree aspect of eBay interaction. Note that Figure 6 actually models two normal-form games in extensive form: the shipping game described in the game theory section above, and a feedback game modeling the decision of whether or not to leave proper feedback, which is defined as accurate feedback for a given trade (i.e. "cooperate"), or improper feedback (i.e. "defect"), which is defined as inaccurate feedback or none at all. In the feedback game, prone to the free rider dilemma because rating takes time to do, improper feedback is taken more specifically to mean rating the other party unfairly, using feedback as blackmail, a phenomenon discussed below, or not leaving feedback at all. I have left out specific payoffs, because they change based on how important agents consider reputation.

* Of course, buyer fraud exists, such as when buyers use carefully crafted schemes to get sellers to ship an item before it is paid for. Tactics include spoofed emails that make it seem like payment has gone through. However, since seller fraud is far more rampant, I focus on it.
Without feedback, the game will result in a no-trade equilibrium. With reliable feedback signaling, however, buyers will theoretically be able to distinguish low-quality from high-quality sellers and deal only with the latter. If sellers expect that they will get positive feedback for a proper delivery and negative feedback for a bad one, then they will consider the future implications of these reputation signals when deciding how to play, thus producing the coordination game posited in the previous chapter that includes reputation effects in players' payoffs. If feedback is not tied to this and loses its power of punishment and reward, the game regresses back into the basic prisoners' dilemma.
Evidence suggests that feedback, while usually fairly applied, is not perfectly correlated to proper or bad delivery. As I will discuss more fully below in the problems section, the vast majority of feedback on eBay is positive, and it often comes in pairs; that is, one transacting party will punish the other for negative feedback by also giving a negative rating, even if the experience was positive.95 Similarly, positive feedback most often begets positive feedback. Such a pattern calls into question the ability of eBay feedback scores to reliably signal cooperation and encourage traders to value the future implications of their reputations.

Partly because of this widespread feedback reciprocity, eBay announced on January 29, 2008 that, beginning in May 2008, it is eliminating sellers' ability to rate buyers negatively. According to eBay, it is making this change because the current system prevents buyers from leaving honest feedback, as they fear retaliation from the sellers if they leave a negative. According to eBay, "this makes it hard for buyers to distinguish between sellers [and their feedback signals] while making bidding or buying decisions. In addition, when buyers receive negative feedback, they reduce their activity in the marketplace, which in turn harms all sellers."96 By improving the accuracy of the reputation system, they claim, transactional frequency will increase; the announcement argues in clear terms that "buyers will bid more and higher since their trust in the Feedback system and the sellers will increase."97 Many sellers argue, however, that eliminating their ability to rate buyers gives an unfair advantage to buyers, who no longer face a credible threat of negative feedback and can potentially extort more goods or services using their own negative feedback as blackmail.98 Nonetheless, eBay hopes this change will improve the reliability of trader signals to
increase trust and volume of trade, demonstrating the company's belief in the importance of reputation in fulfilling their mission for profit. Some pundits, though, see this change as eBay's implicit acknowledgement that trust has broken down:

The theory was that, as in a medieval souk in which everyone knew everyone, everyone on Ebay would know who the crooks were by reading their feedback. Now the company has basically admitted that the cybersouk model does not work: buyers did not tell the truth about sellers, and sellers did not tell the truth about buyers. And in a market where traders lie, the trust that is so central to online commerce cannot flourish.99

Whether the new model can curb this disturbing trend remains to be seen. The lesson is clear: if feedback scores signal false identity information about past cooperation, trust-supported interaction cannot persist in the face of information asymmetry.

On eBay, a higher feedback score gives sellers access to more perks than just increased likelihood of sale. The most notable of these perks is the Power Seller program, which gives sellers special rights and discounts unavailable to others, as well as an icon next to their username advertising their membership in the program. Such an icon is a signal of exceptionally high ratings and communicates to buyers that the seller is especially trustworthy. Requirements for admission to the program include "consistent sales volume, 98% total positive Feedback, eBay marketplace policy compliance, an account in good financial standing, and...detailed seller rating (DSRs) of 4.5 or higher in all four DSRs."100 Sellers in this program enjoy fee discounts on sales, better customer service, increased fraud protection, and access to special message boards. Furthermore, all highly rated sellers get preferential standing in search results, making their auctions more visible to buyers, and spreading their reputa-
Once a seller reaches ten positive feedback points, he or she also gets the ability to institute the "Buy it Now" feature. This option lets sellers forgo the traditional auction format in favor of letting buyers purchase the item for a set price, which can make the item sell faster and give the seller control over the price. New sellers are unable to do this. Finally, as a buyer, positive feedback gives one increased customer support and the right to post more messages on the help forums, which for new buyers are limited. Thus, feedback not only signals trustworthiness, but also gives users expanded rights, the use of which may signal happy customers in the past and a reputation for cooperation.

These extra perks are unsurprising when considered in light of the game theoretic tools introduced in Chapter One. eBay is essentially increasing the cost of using unestablished identities and the payoffs of preserving a good reputation to finesse the payoffs that buyers and sellers face into those modeled by a coordination, rather than prisoners' dilemma, game. The higher prices they may get account for a portion of this, but the purely monetary benefits do not offer enough of an incentive to maintain a good reputation. For a system like eBay's to work, the benefits of building up a good reputation must exceed the costs. If the costs are too high, then members will shed imperfect reputations in favor of new pseudonyms to mask uncooperative identity information: this is the problem of "cheap pseudonyms" that is so common online. Michal Feldman and John Chuang identify the "whitewashing attack, where

* Search position can be viewed as a sign of reputation, because users interpret higher search results as sign of a product that is a closer match to what they want. Highly ranked search results are also more often clicked, leading to a higher probability of a buyer even seeing a certain seller's product and then buying it.
users continuously discard their old identity and acquire a new one to escape the consequences of their bad behavior. As Akerlof explains, this will lead to buyers being unable to distinguish dishonest (low-quality) sellers using this strategy from honest (high-quality) ones, and therefore not bid in the first place. Distrusting newcomers without good reputations is one way to decrease this problem, but eBay clearly realized that this cost was not high enough to discourage trading in for new pseudonyms. The benefits of a positive feedback profile, though not extraordinary, may be enough to motivate sellers to behave and keep the same identity, even in the face of a few negative feedback points.

One of these benefits for the seller, as alluded to above, is the ability to successfully make higher-priced sales. If this is a noticeable effect, then it will be in traders' best interests to signal a cooperative reputation by a high feedback score so that they will be able to take part in profitable transactions rather than driving buyers out of the market. Furthermore, if price effects punish defectors and new, reputation-lacking agents, then they will reduce the benefits of using an endless supply of cheap pseudonyms.

The literature supports the claim that a positive reputation rewards sellers, though not buyers, with higher prices. In addition to being able to refuse to sell to buyers with low feedback scores, the seller is in a favorable position because he or she can wait for the money to come through before releasing the item. From the perspective of buyers, sellers' reputation means everything because they "move" second in deciding whether or not to ship a high-quality good, so it is that measure that I will focus on. David Baron identifies two effects that reputation may have on auctions:
first, it might raise the likelihood of buyers bidding; second, it might raise the price.\textsuperscript{106}

In "The Value of Reputation on eBay: A Controlled Experiment," Paul Resnick, Richard Zeckhauser, and others undertake a review of fifteen empirical studies linking reputation to price. They find that, in general, reputation correlates with a higher selling price, but varies in how much it affects the price, though not every study comes to this conclusion. In summary, they note that "reputation affects the probability of sale as well as price. It also affects the probability of individual bidders entering the auction, and the number of bids. It also affects buyers' subjective assessments of a seller's trustworthiness."\textsuperscript{107} Numerous studies support the conclusion that reputation has a noticeable effect on sellers' ability to sell their goods at higher prices.\textsuperscript{*} This price-effect is especially strong when there is uncertainty about the quality of the goods, a trading situation which requires more trust on the part of the buyer that the seller has accurately described the goods on the auction page.

Given the evidence that positive feedback warrants higher prices, it is no wonder that sellers closely guard their reputations and put significant time into building them up. In terms of the game theoretic model posited earlier, this ensures that reputation effects are seen as a meaningful punishment and reward that encourages traders to consider future interactions rather than approaching each auction as a single-shot prisoners' dilemma. Linking payoffs in this way, the feedback mechanism creates a coordination super-game. Feedback is a signal, though instances of fraud that will be discussed reveal that it may not always be a reliable one, that a seller is trustworthy

\* For more information, see in particular Resnick et al., 2006; Melnik and Alm, 2005; Resnick and Zeckhauser, 2002; Houser and Wooders, 2006.
and will deliver the goods as they are described. While it works more effectively than one might expect to support faceless traders buying goods they have never touched, the system is not without its problems. eBay has attempted to address some of these problems, as discussed above, with the new feedback system changes, but it remains to be seen how effective they will be.

2. Problems

The eBay feedback system, though impressive, does have problems creating the trust that supports the site's purpose. For the system to be effective, reputation must reliably predict future cooperation, but due to free-riding, feedback extortion, and fraud, this is not always the case. In other words, problems such as these may make trader signals unreliable and lead to the persistence of information asymmetries and their deleterious effects on trade. In this section, I will discuss these stumbling blocks.

Free-riding, though a problem, is not as much of an issue as game theory would predict. In a prisoners' dilemma, the rational self-interested choice is to not give feedback, because it takes time and effort. Once a transaction is completed, there is no personal gain (notwithstanding the pleasure of recognizing a job well done or punishing a dishonest transactor) in taking the time to rate the other trading partner for other members' benefit. Some have expressed surprise that the percentage of members—over half—that provide feedback is so high. Others have tried to explain the phenomenon by pointing to the bonds that are forged between buyers and sellers in the exchange of messages during the auction, or by positing "some quasi-civic
One final explanation, which I believe is the most satisfying, links to the next problem on eBay: feedback extortion and reciprocity is a major phenomenon on the site. The hope of feedback reciprocity (traders "repaying" positive feedback with positive feedback) adds a benefit to leaving feedback, making it the more economically rational choice.

Feedback extortion is a problem with the eBay reputation system, though it remains to be seen whether this will shift as a result of the policy change that sellers can only leave positive feedback for buyers. This extortion phenomenon occurs when traders withhold positive feedback until they get it from the other party, or when they use the threat of negative feedback to demand more goods or a lower price than agreed upon. Resnick and Zeckhauser note that "there [is] a high correlation between buyer and seller feedback, suggesting that the players reciprocate and retaliate." They find that the seller rates the buyer positively 99.8% of the time when the buyer rates the seller positively, but "only 39.3% of the time when the buyer is neutral or negative," and that similar trends hold vice-versa. The authors call this the "high courtesy equilibrium," though it may be a discourtesy to other users, who rely on ratings in the decision of whether to enter a bid. They also find that an unusually high percentage of auctions result in positive feedback, which is surprising given the reports of fraud on the site. This is likely due to the fear of retaliation.

This fear leads to dishonest reporting and therefore reduces the ability of the system to reliably predict future cooperation and encourage trust. eBay cited this problem in its explanation of the recent changes: "The current system prevents buyers from leaving honest Feedback as they fear retaliation from the sellers if they leave a
negative. Again, it remains to be seen what will emerge as the new equilibrium now that sellers can only leave positive, or no, feedback for buyers. Buyers may be more likely to leave negative comments when they know they cannot be hurt as much in return. This could lead to more honest reporting, or the reverse, though this seems less likely. It does not solve the problem of more general extortion, such as buyers demanding special services or more of a product by using negative feedback as a threat. This works especially well against established sellers, who know that one bad mark can tarnish their recent feedback scores and lead to lower priced sales or removal from the Power Seller program. Without the recourse of rating the buyer badly, they lose this credible threat as a balancing agent against dishonest buyers. Nonetheless, the change may curb problems of feedback reciprocation, in which traders leave feedback for purposes other than honest rating, but it might also increase the free rider problem as buyers see less of a reason to leave feedback without the threat of negative reciprocation.

Outright fraud—especially that based on the dishonest inflation of one's reputation—and the knowledge that this happens, is a third obstacle to reputation fulfilling its role of supporting trust that ensures long-term cooperation. The fundamental aspect of online interaction is that one usually does not know with whom one is dealing. As has been discussed, this can be a positive, by connecting people who might not otherwise connect, but it can also be a negative, especially when real money is at stake. For this reason, trust is hampered when participants in a trading environment such as
eBay do not know whether the feedback behind a pseudonym is legitimate. Dishonest sellers "game," or take advantage of, the system by exploiting its flaws. One way to do this is by building positive feedback by selling one cent items, such as e-books, that are costless to copy and distribute. After a sale of one of these, buyers leave positive feedback, and a few hundred of these leads to a high seller rating. According to John Morgan, in an interview with *The Wall Street Journal*, "If I am a bad seller and I engage in a one-penny transaction, that's one feedback point. And that is just as good as a good seller that engages in a $1,000 transaction." After building up a good reputation with these types of transactions, a seller can offer a bigger ticket item, take payment, and then disappear to start the scam again under a new pseudonym. Knowing that this happens can activate the risk aversion of buyers and drive them to the defection equilibrium in the face of uncertain information about whether the seller is going to cooperate or defect. This progression results in the no-trade equilibrium modeled earlier.

Another method criminals use is to simply hijack accounts and use them in conjunction with stolen credit cards, as a buyer, or to take payment for nonexistent goods, as a seller. Additionally, a quick search on Google reviews numerous instructional pages on how to use "phishing" scams to create bogus websites that look like eBay and harvest account information, and javascript codes that mimic the look of a real feedback score. These types of scams target the necessary condition for cooperation in an environment where reputation stands in for partners' repeated inter-

* Post-sale fraud, such as money-order or credit card scams, is also common, but has less to do with reputation and is most often committed in tandem with the use of a false feedback profile.
actions: when one cannot identify one's playing partner, the basis for choosing a co-
operative strategy disappears and defection becomes the best option. Stolen accounts
make buyers worry that feedback is a reliable signal of the trustworthiness of an ac-
count, but not of the person using it at the present time. In other words, if reputation
is known not to accurately reflect the true identity of the user behind the pseudonym
(as a person who has cooperated or defected in the past), then the trust it supports will
collapse. This results in the information asymmetry problem–only the lying seller
knows that the current game is his or her last. Furthermore, when defectors are able
to hide behind cooperative pseudonyms, these users will drive out actual cooperators,
because feedback will become known as an unreliable signal of future cooperation.
The fact that eBay is still around suggests that this point has not been reached, but the
reputation system clearly has flaws and is open to fraud.

An overarching problem that exacerbates these three issues is that people do
not make enough use of the information available to them. One of the distinct bene-
fits of Internet-based reputation systems is the low cost of spreading information,
which increases the number of interactions that are possible. But this information is
not always efficiently used. As one article puts it in trying to explain why negative
feedback did not affect the probability of sale for new buyers, "most buyers simply
did not bother to click through to look at the detailed feedback, and instead merely re-
lied on the overall score...that eBay displayed as part of the lot listing."118 If buyers
peruse the text comments as well, in order to see what items the feedback is for, they
are able to see whether this seller has sold similar items in the past, or if the item they
are considering purchasing is an outlier, a signal that something might be awry. Fur-
thermore, email communication with the seller may reveal inconsistencies in identity, such as when the seller displays poor knowledge of the product. The simplicity of eBay's feedback system no doubt contributed to the site's popularity as it grew, but this simplicity also hinders the mission it is supposed to fulfill: creating trust and predicting future cooperation to enable trading. If buyers do not take the time to use the signals at their disposal and consider carefully with whom they place a bid, misrepresented identity and extortion are more likely to occur.

3. Conclusions

eBay was one of the first consciously designed online reputation systems that expanded for use on a large scale. It is also one of the simplest, at least on the surface, and considered important by its users, making it an ideal test case with which to discuss online reputation. eBay's mission, and that of its users, is clear: profit. Sellers post items to make money, buyers bid on them hoping for low prices or unique finds, and eBay itself profits from each sale. It is thus in the company's best interests to enable trust between buyers and sellers, so more items will be sold, which leads to higher profits. Early on in eBay's history, feedback mechanisms emerged as a necessary tool to fulfill this mission. As the number of users increased, growing pains affected the system and modifying the feedback mechanism was necessary.

Overall, the feedback system in place effectively supports eBay's mission. But it is far from perfect, and the boundaries of such a simple system are beginning to be felt. Though eBay does not release specific statistics, anecdotal evidence suggests that many users have experienced some sort of fraud or identity misrepresentation on
the site. High volume customers can absorb these loses, but they may be enough to drive away the occasional user, to whom it may be worth paying a slightly higher price to buy the same good from a more established seller (such as Amazon.com or a brick-and-mortar store) to avoid the risks of using eBay. Of course, some items cannot be bought from these other stores, leaving eBay as one of the only options for niche markets, such as that first broken laser pointer sold. The benefits of eBay are certainly impressive—wide selection, high trade volume, and lower prices—lending support to Resnick and Zechauser's argument that "perhaps a fairly effective reputation system is good enough."

eBay demonstrates that even a relatively simple rating system is far more complex when looked at under the hood. Such examination also reveals many of the problems with feedback mechanisms that attempt to enable reputation reporting. People will always find a way to take advantage of the system. This is a particular danger in a closed system like eBay's that only makes use of reputation information from within eBay. So while eBay illustrates the benefits of a reputation system, it also suggests the need for enlarging the "coalition" to improve signals and reduce information asymmetry. A seller might have a great reputation on eBay, but a horrible credit rating filled with a history of scams. When the two exist in different places, without the possibility of connection, buyers lose out on important information they could use to make their strategy decisions. In short, eBay is just one reputational context about a person, but one that stands to benefit from reciprocal sharing with other forums. The information it already has would also be valuable to other contexts that hope to support trust among their users. eBay's system, however, works remarkably well, partic-
ularly in its ability to support trust via a reputation mechanism that overcomes many of the theoretical risks of one-shot interactions by linking present to future payoffs.
III. WORLD OF WARCRAFT

Blizzard Entertainment released World of Warcraft (WoW) in 2004 as a massively multiplayer online role-playing game (MMORPG) based on the popular computer strategy game Warcraft. In WoW, players control characters who go on quests, and play takes place in real-time with other players who are connected over the Internet. WoW is not a website, but rather a program that sends information over the Internet so that players can interact with one another. Rather than play against a computer, users compete against and work with others. The game is social, unlike the solitary video games of the past. In this capacity, MMORPGs are generally considered to be the descendants of multi-user domains (MUDs), which were multiplayer "virtual worlds" based in text.\(^\text{120}\) Other popular MMORPGs exist (such as Everquest and Dark Age of Camelot),\(^\text{121}\) but none comes close to the popularity of WoW, which recently reached ten million active subscribers and owns over half the market-share of MMORPGs.\(^\text{122}\)

Before moving on, it is necessary to gain familiarity with the layout and vocabulary of the game. WoW play takes place in Azeroth, a fantastical virtual world

* William Bainbridge describes a virtual world as an "electronic environment" that "mimics complex physical spaces" and enables interaction between animated characters (who represent their human player and are often called "avatars") and objects (Bainbridge, 2007, p. 472). Edward Castronova adds that virtual worlds are "crafted places inside computers," raising the point that the game designers have a large measure of control over how players interact with one another (Castronova, 2006, p. 4), but these players do also decide how to use the tools designers give them. See also Lessig, 2006 for his famous "code is law" argument and its implications for how players interact.
reminiscent of Tolkein. Dragons, trolls, and other monsters present obstacles that players must overcome in order to succeed in the game. When starting out, players get to design their character, selecting from different preset races, such as human and elf, and classes, such as warrior and priest. Depending on one's race and class, one has access to different skills, such as the ability to cast spells (needed in battle to subdue one's opponent or heal wounded players) and develop trade skills such as blacksmithing or tailoring. These skills allow players to make money by selling goods or trading with other players. A player can have multiple characters in the game, but may play only one of them at a time.

The heart of WoW is quests. When players complete these missions, they receive prizes such as items, skills-training, or gold (the in-game money). The types of quests range from killing a certain number of stray wolves, a solo mission, to fighting a dragon in a type of mission called a "raid," which requires team cooperation of up to thirty players. For every enemy killed, a character gets a certain amount of "experience points" (XP), and the number of XP accumulated determines a character's level, which is currently capped at 70. "Leveling up," as WoW players call improving one's level, has rewards beyond pride: players gain access to new items (such as swords and armor) and the ability to learn new skills. Leveling up is thus a critical activity if players wish to access more challenging parts of the game where a lower level character would quickly perish.

Players must bond together to complete many of the game's quests and acquire certain items. Grouping is done in two ways. First, players can group temporarily into a "party" to tackle a quest or series of quests. Membership in a party
continues until a player voluntarily leaves the group, gets kicked out by the party leader, or signs off. Party membership means that all accumulated XP is split between the players. What to do with "loot" (items dropped by defeated foes), however, is left to the discretion of the players in the party, who often decide how to split up loot before setting out to quest.* Because grouping in a party is often temporary, players must evaluate other players' potential value as group members, in terms of both game skills and a fair ethic in splitting up loot. Due to the transitory nature of parties, members do not have the power to punish defection (defined below) in future interactions.

The second type of group is a guild. Guilds are groups of players that form to give one another a more permanent "support system." Unlike temporary parties, one's membership in a guild can last between playing sessions; even when one logs off, one's guild affiliation persists. At any time, however, a player can be kicked out of a guild by a guild manager, or they can choose to leave. Guilds have special chat channels reserved for members of the guild, the size of which varies from a few characters to over one hundred characters. The level of seriousness with which guilds approach the game also varies: some consist primarily of friends who enjoy playing together, while others have complex recruiting procedures aimed at ensuring that every player in the guild is skilled and takes the game seriously. Because characters in guilds regularly quest together, similar evaluations to those used in temporary par-

---

* WoW does provide some in-game systems for allocating loot, such as random distribution, but because players often decide by other means (sometimes based on how much a player "needs" a certain item), questions of how to split it up are important enough to warrant consideration in this chapter, especially of the potential these activities present for cooperation and defection.
ties are needed to ensure that players have a certain level of skill, will respect loot sharing arrangements, and will help fellow members out when needed. Some guilds employ complex tracking systems for this purpose, but many rely on informal reputation monitoring not so different from Dellarocas' word-of-mouth systems. The fact that players often stay in guilds for long periods of time makes it is easier to punish improper play in future interactions. Punishments include withholding loot or banishing players from the guild, which affects their ability to find future playing partners.

The importance of grouping suggests the roles of trust and reputation in social interaction within WoW. Grouping with other players, whether in a party or guild, requires trust that other members have the skill that one expects them to and will fulfill their obligations to the group. WoW groups lack the luxury of legal contracts to guarantee cooperation, so players must interpret signals to decide whether their group members will cooperate in such ways as playing an agreed upon length of time (until a certain quest or quests are completed or until each player gets the item he or she needs) and splitting up the loot in agreed upon ways rather than trying to get as much for oneself and then leaving. Grabbing all the loot (defection) is known as "ninja looting" and can lead to negative reputation effects, which may affect a player in ways discussed below. Another type of defection is misrepresenting one's skill, which can lead to one's own or teammates' death in battle. Planning to steal loot,

---

* One of the unique aspects of WoW is that these trusting relationships may extend beyond the game, as is evidenced by numerous relationships that escape the game's boundaries, and which suggests a "community" status inherent in certain aspects of WoW. These considerations, however, are largely beyond the scope of this paper. For more information, see in particular Yee, 2007.
leaving before teammates complete the quest, and misrepresenting one's facility with the game are all examples of an information asymmetry problem, because there is no way for other players to know the defector's intention.

If this problem is widespread, it threatens to break down social interaction as players lose the ability to distinguish between cooperative and uncooperative players in a series of single-shot interactions. Thus, reputation is necessary to signal skill and cooperative play, lower risk aversion, and provide a way to punish defectors that connects present to future payoffs. Edward Castranova also notes the role of "reputational capital" that arises when items are scarce, as they are in WoW: "Since everything is not free, and since you need other people to get the things you want, you had better behave....Reputation and norms [of cooperation] are often more powerful than law." If a character develops a bad reputation, others may not group with him or her, making it impossible to complete the game. Networks among players and ample central public guarantee fast communication. News and gossip travel fast in this world, so if one gets a bad rap (such as by stealing items or not showing up for a planned quest), other important players have a way to find out quickly. In this chapter, I examine whether or not reputation actually fulfills this signaling role.

WoW's popularity is astounding. According to Nicolas Ducheneaut, Nicholas Yee, Eric Nickell, and Robert Moore, who measured usage statistics eight months after the game's release, the average time of play for a level 60 character [the highest level attainable at the time] was 15.5 days. I quote these statistics to underscore the role that WoW plays in some people's lives. Yee notes that players put a lot of stock into advancing levels and in-game socializing, often as an escape from offline prob-
lems. One doctor estimated, based on this work, that 40% of WoW players are addicted to the game.\textsuperscript{126} To many players, the reputation and recognition that comes from having a powerful character is a meaningful and deep source of pride. So while it may seem as though reputation might not matter in a setting such as WoW, because it is "just a game" where nothing real is at stake, such a view does not account for the motivations of hardcore gamers who place more importance on the game than they do on their offline lives, often finding themselves playing (and socializing within the game) for hours upon hours a day.\textsuperscript{127} Because success in the game is so important to many players, their reputations as skillful and cooperative players becomes a motivation in and of itself, further intensifying the power of reputation to reward and punish in such a way that makes players consider the future effects of their actions.

Understanding the purpose of WoW is also necessary to understanding why trust and reputation are important. Blizzard, the company behind the game, wants as many people as possible to buy their software and keep playing WoW. The game software costs only twenty dollars, but requires a fifteen dollar a month subscription, so the longer players stay engaged, the more money Blizzard makes. Improving a character keeps players engaged for a while, but it is the pleasures of reputation and social interaction that keeps them around. The purpose of the game, at its simplest, is to attain the highest level and all the best items, but other motivations account for players' participation as well. Once an impressive reputation (and therefore bragging rights) are established within the game, it is more difficult to cancel membership because all of one's accomplishments are erased. Enabling a reputation system that en-
ables players to signal these accomplishments (identity information) is thus critical to getting them to play more.

Admittedly, the purpose of Wow is less clear-cut than that of eBay or Facebook. Nonetheless, it is there just below the surface, and certainly surpasses the simple entertainment value that many critics ascribe to it. It remains to be seen, however, how well the reputation mechanisms in place support these purposes. Understanding this will not just help us understand reputation systems in WoW and other games, but also the feasibility of "performance-based" reputation systems, like that which is in effect on WoW and that enables reliable signals. Performance-based systems, as will be seen below, rely on speech and action to spread reputation and require interpretation on the part of the identity requester. When a player slays a dragon or brags about a little-known place in the game, this can be viewed as a performance that broadcasts reputation. This makes the reputation system not unlike reputation in everyday life, minus the body language people use to make judgments in person. The theory predicts that such a system, even with its required interpretation, would still effectively overcome information asymmetry with signals that support social interaction. It is to this system that I now turn in order to evaluate this claim.

* For more information on player motivation, see Yee, 2007.
1. World of Warcraft's Reputation System

World of Warcraft presents an interesting case study for examining reputation's effect on trust and purpose, because its participants take the events that happen within its virtual boundaries seriously, and interaction (mostly in the form of chat) happens without the rules that govern on eBay. Players can talk about what they want and do what they want, as long as they have the skills. The game actually operates using two reputation systems: one for player characters (PCs) and one for computer-controlled
characters, but I will focus here on the reputation system in place for player characters, because this is the system that matters most for social interaction.*

Signaling one's reputation to player characters is the basis of social interaction in WoW, because it is how players make grouping decisions. Because players' opinions of what constitutes an experienced and knowledgeable player are largely subjective, explicit measures do not work well, unlike on eBay, where a seller who ships fast and represents the item fairly is universally considered worthy of a high rating. Before moving on to the items- and performance-based reputation system in operation among players on WoW, however, it is useful to note that there is one form of explicit reputation (that type which can be boiled down to a simple number): one's level. Beside every character's name in WoW is a number representing that character's level, currently ranging from 1-70. In simple terms, a higher level marking corresponds to a more skillful character. It means, in theory, that one has completed the increasingly difficult quests in the game and therefore is skilled. The truth is more complicated than this, because people can buy pre-leveled characters online, but I will discuss this more in the problems section. The problems with level rank make it a reliable signal of the character's abilities in the game, but only a conventional signal of the skill of the player using the character. Level ranking does not signal any information about prospects for cooperative play.

This leaves items, speech, and action as possible signals of reputation. I will start with items, often considered one of the most significant measures of reputation

* NPC reputation is exclusively tracked by the game. For more information, see Ziebart, 2008; BlizzardEntertainment; Castronova, 2006, pp. 305, note 4.
corresponding to mastery of game, but one that has significant problems and can therefore only act as a conventional (forgeable), rather than reliable (handicap, and therefore trustworthy) signal of skill, just like a character's level ranking. As I mentioned, in WoW, defeating enemies and getting in-game money leads to the acquisition of various prizes, like armor, swords, or new spells.* WoW is filled with colorful, intricately designed items with fantastical names such as "Bulwark of the Ancient Kings" and "Blade of Vaulted Secrets." Most items are restricted by level. For example, mounts, rideable animals that allow a player to travel significantly faster, are only available to players at level 40 and above. As such, mounts are considered "a significant social status marker of being a 'high level' character.\textsuperscript{128} Being high-level does not guarantee a character access to rarer items, which still must be acquired from "drops" in difficult quests, referring to the loot dropped by a vanquished enemy. Therefore, certain items signal that a character has been playing an especially long time and has been able to complete enough quests or get enough gold to acquire them, but the signal can be faked by purchasing characters or items. Like level, items do not signal information about cooperation.

Displaying these special items is a way to show others one's skill, which is a motivation in and of itself and supports the game's purpose of giving successful players an ego boost. In this way, items contribute to one's reputation, creating a model for punishment and reward. The reputation effects of earned possessions is a central feature of the MMORP environment. As one blogger writes, "whatever you possess

\textsuperscript{*} Though not an "item" as such, many spells are only available to skilled players who have spent time in the game winning quests and building up enough money to pay for training.
in that world [an MMORPG], your friends can see it....[and] to obtain a rare item or to be an advanced player gives you bragging rights..." The power to display one's accomplishments like this is crucial to players' enjoyment of the game. In fact, Yee finds that, at a statistically significant level, players are motivated by "acquiring rare items that most players will never have," "accumulating resources, items or money," and "becoming powerful." These are motivations that gain their value through public performance. In solitude, leveling-up and getting a rare item would bring most players less satisfaction, because there would not be others to see them and be impressed. In other words, "players could not compete [especially in reputation building] if there were no one to play with." Ducheneaut, Yee, Nickell, and Moore explain this idea well in an extended passage:

To begin, it is important to mention that MMORPGs are in essence reputation games—an avatar wearing powerful items, for instance, is essential to the construction of a player's identity. It broadcasts the player's status to others and rewards him or her with a sense of achievement...Put differently it is not "the people that are addictive" but rather, "it's the image of myself I get from other people." Displaying one's items bolsters reputation by signaling skills and playing power, which in turn boosts one's ego. In short, the reputation is the reward.

Players signal their reputation by "peacocking," displaying their best items and wardrobe for the reward of the admiration of others and to demonstrate one's gaming skill. As I addressed in the section on animal signaling in Chapter One, for a signal to be reliable, biologists have explained, it must waste the resource it is supposed to indicate. In the context of WoW, rare items waste time and in-game gold, thus signaling that a player has accumulated hundreds of hours of game play and
gained the skill that this play brings. This assumes that the items were not bought from online character and items dealers, a phenomenon discussed below that would make these signals only conventional. For now, however, I will assume that all character traits are legitimately acquired. One study reports that "it is not uncommon to see level 60 avatars, wearing powerful sets of armor and weapons, simply left standing by their players in front of the auction house for everyone to admire." One player, posting on a WoW message board, bragged about his or her character's appearance and the compliments it garnered: "I especially like to customize my own casual looking outfit. and I know it looks it too, cause random people walk up to me [in the game] and say so." This sort of behavior works as a signal of reputation, because many players are closely familiar with the value of different items and clothing combinations, and it makes players who have worked hard to succeed in the game feel good. Showing off one's items can also be a strategic move. Many quests require grouping with others who possess certain skills or items that will make it easier to overcome an opponent. Displaying these skills and items is a reliable signal that one will be able to bring them to battle and leads to more offers of grouping, which in turn leads to the experience points needed for advancement. Note that this display does not signal the likelihood of one's cooperative behavior. Signaling is therefore not so different from the animal world, where wasting a resource signals an excess of the resource and makes it easier to broadcast reputation and evaluate others, necessary tools in a game where grouping is needed to excel.

Unlike on eBay, a good reputation does not lead to material gains. Even though reputation helps in advancing within the game by signaling to others that one's
character would make a good member of a group or guild, most of the benefits of reputation come in how it makes players feel. This acts as an incentive to keep up cooperation with others when in a group. The feeling one gets from being labeled a "dirty player" may be enough of a punishment to reduce this type of behavior. Furthermore, being labeled a defector will negatively impact one's ability to group and advance in the future, thereby transforming multiple prisoners' dilemmas into one coordination game. Reputation is also an incentive to continue playing the game. In order to get more rare items and the reputation boost that comes with them, one must continue playing and renew membership. This highlights how reputation serves Blizzard's business interests.

Player profiles are automatically posted on a public website called The Armory, which lists characters and their items, skills, and attributes in a searchable database. Most information is accessible by anyone, though some requires a WoW account. Blizzard added this feature in 2007, most likely in recognition that it would increase the reputational reward players feel, because even non-players can look at how accomplished they are in the game. The Armory also begins to make players' reputations accessible outside of the game, though note that it only promulgates conventional signals (such as items and character level) that are susceptible to the problems discussed below. Still, The Armory is an excellent example of a company realizing the importance of reputation, especially one that is accessible from beyond the confines of the service, for grouping decisions and player ego and introducing a tool to spread this reputation more publicly.
Players also signal skill by the things they say (speech) and the actions they take. Speech and action, in that order, are more reliable signals than items, because they cannot be bought. In other words, while a player can buy a pre-created level 70 character with lots of gold and the best items, he or she will not necessarily know how to use these skills. One WoW commentator writes, on a popular blog, about selling his characters that, "The level of competence required from endgame [quests] is very high, and those who don't measure up are quickly blackballed, so I feel a certain remorse in taking these people's money [in exchange for a high level character] and throwing them to the wolves." Though the ability to cast certain spells and use certain weapons and armor brings power, not knowing when or how to use them signals weakness and an unfamiliarity with the game. Understanding this, players engage in activities such as skill competitions, arguing over the best strategies for different quests, getting into debates that display their "familiarity with game/characters," and showing off "tricks to complete tasks in the game." Familiarity with the game's culture also signals veteran status. There are different humorous tricks that experienced players will pick up, such as dancing on in-world mailboxes or wearing items at inappropriate, but clever times. One article describes a character wearing a deep-dive helmet in an auction house. These activities may seem pointless, but they signal a sense of humor (which people value in teammates) and knowledge of WoW tricks that could only have been gained by playing the game for a long time. These signals make a difference for whether others will trust a character enough to group with it. In addition, a player who does not know the jargon of the game, even if his or her character is high-level and possesses rare items, will not pass as an experienced
player. Sometimes though, players must engage in costly evaluation by actually playing with another player to make sure he or she is skillful and trustworthy. Thus, speech and action are reliable (the latter more so) ways to show others game knowledge and an important aspect in the evaluation of other players, especially when deciding with whom to group.

The guild system within WoW is an important facet of the game that both acts as a signal of reputation and requires reputation to function smoothly. A guild is a community within the game, one that enables social capital by giving players a social network through which to ask questions and attain hard-to-find objects. Within the game, if a player belongs to a guild, the guild's name is displayed under the character's name. Membership in top guilds, such as Nihilum and Aurora, garner respect from other players. It also signals skill, as "admission to these prestigious social groups often requires going through a 'trial period,' as well as being sponsored by one of the members." In the "Life and Death of Online Gaming Communities," Ducheneaut, Yee, Nickell, and Moore explain that "Being a member of an "elite" or "uber" guild, renowned for its ability to tackle the hardest challenges, is therefore a badge of honor." This badge is a more trustworthy signal than mere level, because it is unlikely, at least within the higher ranked guilds, that a player could sell his character without someone else in the guild knowing about it, which would lead to the character being kicked out of the guild. Because players sometimes quest with characters

* Just as how an offline interviewer might think about how the person being interviewed would be to work with, many guilds look at other things besides skills, such as personality and attitude, to judge potential members.
outside of their guild, a prestigious guild badge also acts as a reputational signal for others to use.

In addition to guild membership acting as a signal, reputation is critical to the functioning of guilds. Tony Manninen and Tomi Kujanpaa explain that "in guild activities, concepts such as trust and reputation become essential as part of the player image. Some of the guilds require a certain amount of playing hours or certain percentage of attendance in guild activities, such as meetings and raids." If guild responsibilities are not lived up to, players may get kicked out. Due to the limited in-game tools (consisting of a guild-roster and guild-chat) to track these activities, well-organized guilds use outside services to schedule quests and track participation, which can be important in deciding who gets certain items. Most of these rely on an in-guild currency called dragon kill points (DKP) that players get for participating in raids and other guild activities. DKP can then be exchanged for access to certain dropped items. DKP thus stands in for word-of-mouth systems by statistically tracking players, and the amount of DKP a player builds up is a signal of reputation. Guilds rely on this reputational reward to keep players interested and to fairly divide the spoils of war, which makes having a reputation as a skilled, cooperative, and engaged player a valuable reward within the game.

Participation in a guild can sometimes manifest as a series of prisoners' dilemmas, especially the free rider problem. If one just shows up for a quest, but does not help out much, perhaps by leaving the computer to do something else, there is a chance others will not notice and one will still get the benefits of participation, such as experience points and loot. Of course, if everyone did this, no quests would be
completed. In addition, a player can just leave a guild once his or her objectives, such as getting a certain item, have been completed.\textsuperscript{144} Non-participation can jeopardize other characters. Manninen and Kujanpaa explain that, "Since the death of a game character often results in the loss of experience points and other virtual assets, the players generally feel the need to trust in each other."\textsuperscript{145} As the theory predicts, more information about a player will lower risk aversion and make players comfortable questing together. Reputation, such as that gained by assigning DKP, helps overcome this problem by identifying players who are not "pulling their weight" and should therefore be kicked out of the guild. However, this is easier said than done, and the problem of guild management will be discussed further in the problems section. Nonetheless, defectors face a credible punishment of being kicked out of the guild by guild-mates, who do not want to play with free riders and may take broken trust personally. Players resist grouping with characters who have defected in the past and the losses that would result from a repeat of this defection. Since guild free riders, and other types of defectors, may find it harder to find others to group with, future consequences will impact their immediate playing decisions.

Reputation in WoW may be, as was discussed in relation to eBay, "good enough" to support many in-game activities where imperfect reputation does not significantly harm interaction. Especially as a new player, interaction with other characters is critical to learning about the game, but one has to know whom to ask. On various chat channels, players often ask questions such as "how do I get to the badlands [a place in the game]" or "what's a good strategy to beat the traitor's destruction quest?" Bonnie Nardi and Justin Harris explain that "because characters are labeled
with their name and guild, and mousing over the character reveals class and level, players have enough information to make appropriate requests.”146 If the person behind the character is not as experienced as he or she seems, only the typing of a short sentence has been wasted, but for other larger tasks, such as deciding whether to group together on a tough quest, imperfect reputation can be a problem. This sort of problem is often mitigated in guilds, because if one trusts his or her guild, it follows that others in the group can be trusted for advice and partnership. Thus, reputation is stronger within a guild, but also operates well enough within the general game to support its purpose of getting users to play longer and enjoy the ego-boost they get from their reputations. As such, reputation gains the power of reward and punishment that is needed to shift players' payoffs in such a way that makes them value the future implications of the reputation effects from a social interaction. Reputation thus supports, to some extent, the trust needed by players to enter into social interactions without the fear that the other party will defect.

2. Problems

The WoW reputation mechanism, as we have seen, is much less explicit than eBay's. Although level number and guild membership act as explicit signals, most of a character's reputation is signaled in far hazier ways. Items and spells that are only available to experienced characters indicate that a character has been part of the game for a long time, but because they can be bought on the Internet, this seemingly reliable signal turns into a conventional one. When evaluating other characters, players are left to form their own opinions, based mostly on observations of how others "car-
ry" themselves within the game. Their knowledge of game information, slang usage, and dueling skills are all signals of questing ability, and therefore of being a good partner. Such a "mechanism" can hardly be considered a reputation system at all, as it is more like the way that people evaluate others offline, although this might be part of its appeal: the challenge of recruiting and keeping together a group or guild is part of the fun of the game. Leaving much of the reputation up to interpretation lets players figure out signals, which in turn makes the information they use to inform their playing strategies more nuanced.

In this section, I will discuss some of the problems with reputation in WoW, particularly noting whether reputation supports trust and the game's purpose. From Blizzard's point of view, recall that this is to get as many people to play as much as possible; from the player's point of view, it is to level higher and gain power within the game, especially in such a way that others see this happen. The primary problem, I submit, is that reputation in WoW signals skill (though even this has problems) but not cooperative potential, and that there are few credible punishments for a player who defects.

First, as has been suggested, people buy characters and the reputation they signal. This is a large business, estimated at $1.8 billion across all online games, and it takes many forms.\textsuperscript{147} One of the most popular is what is colloquially referred to as "Chinese gold farming."\textsuperscript{148} This is the practice of mining gold in WoW (and other MMORPGs) and selling it to mostly European and American customers for real money. On sites such as Power-Gamers.net, one can buy 100 gold units (a significant amount) for less than $10, and then use this gold to buy in-game items at auction.
from other players. A search on Google turns up hundreds of other, similar sites. These sites usually offer power-leveling (giving one's character to someone else to play for a certain time) and full character sale services. Many sites act as clearing-houses for these character auctions, and characters have sold for thousands of dollars, with a level 70 going for around $500, depending on the items and skills it comes with. These services give players a way to instantly become a level 70 with the rarest items but without the playing skill necessary to do this on their own, making level and items only a conventional signal. This in turn undermines trust and makes players more suspicious of one another.

The underlying problem is that while assessing player skill is feasible, there is no formal way to assess whether a player is likely to cooperate by being a "fair" player when divvying up loot and remaining in a group until an agreed upon time. Level and item possession may not be a reliable way to judge skill, but since action and speech are, this evaluation is possible. To judge another player's cooperative potential, however, players rely on word-of-mouth systems and the claims of the player under review. Word-of-mouth systems rely on chatting within the game and even posting defecting players' character names on online message boards, but because there are so many players in WoW on many servers, player information may not travel as

* Interestingly, until recently eBay acted as the primary forum for this activity. It banned the sale of virtual items in early 2007, referencing, of all things, the problem with completing the transactions and tracking sellers' reputations (Dibbel, 2006; Anderson, 2007). Also, purchased characters may come with a good or bad reputation that the buyer is unaware of. Furthermore, skilled players may level-up a few characters, keep one, and sell the rest, making it difficult to tell if a player is using his primary character or not and resulting in reputation signals that do not always accurately broadcast identity (playing history).
Players planning to defect may try to convince others that they are cooperators by making friendly small-talk before a quest. One site even suggests that loot stealers "make the players feel comfortable with [their] character[s] by throwing a joke or two [and] get the group laughing" in order to build false trust, and then steal loot. Such behavior may lead one to be labeled a defector by any player who witnesses it, but because it is relatively simple to "disappear" into the world and start up the scam with other players, this reputational label is not a credible punishment. This possibility leads to an information asymmetry between players when they are deciding whether to group up. Thus, there is the potential for a defector to send false signals and defect in a series of single-shot prisoners' dilemmas that do not link together into a coordination super-game. Observational evidence suggests that the game has not devolved to this level, however, and that experienced players are adept at assessing potential teammates.

Reputation within guilds is also difficult to track and leads to cooperative inefficiencies, but guilds' power of punishment makes reputation more powerful in these subsets of players. Because the game only provides a guild roster and chat channel, organizing members and keeping track of activities is laborious. To solve this problem, many guilds turn to outside management tools such as message boards, DKP trackers, and voice-over-internet-protocol (Internet phone) programs such as Vent. Ducheneaut, Yee, Nickell, and Moore discuss how making guild activity (such as ranking and recently completed quests) more public, such as in The Armory, could

* There is even a game add-on called "Black List" that lets one add offending players to one's black-list, but because this information is not shared, it is not as useful as it could be.
lead to better signals.\textsuperscript{152} In short, they argue, the more public reputation is, the more powerful it becomes as a punishment to defectors. Easily accessible guild and player rankings would accomplish this and reduce guild death, a widespread occurrence.\textsuperscript{153} It could also improve the performative aspect of the game. Tracking guild activity participation would add another explicit reputation mechanism to WoW’s arsenal, and bring with it the benefits discussed in the theory of linking reputation to trust. Due to their small size, however, guild members can usually pass along word-of-mouth information that gives a basic idea of others’ cooperation, and defection is punishable by being exiled from the guild. This credible punishment makes reputation more important within guilds than in the larger game and makes guild play a series of social interactions best modeled by a coordination game.

These problems suggest the benefits of expanding the "coalition" among which reputation information is shared. Reputation may work better in guilds because there is a fixed group of players who are likely familiar with one another's behavior. Defectors are more quickly identified and punished within guilds than in the general game, where they can prey on a new group with relative ease. All players, even those not associated with a guild, would benefit from this type of information about cooperators and defectors to reduce the asymmetry with which they enter relationships. Increasing the size of the coalition would require a more effective central repository of player reputation information that mimics the information gleaned by word-of-mouth, because this latter system does not operate effectively among the large numbers of players in WoW who may never encounter one another. As Greif’s medieval trading example suggested, sharing this information among a bigger group
of people would enable more accurate reputational judgments about players and give reputation more power to punish and reward.

3. Conclusions

A mixture of an explicit and performance-based reputation system governs interactions among players in WoW. Such a complex system makes sense when it is understood that WoW is an immersive world that mirrors offline life in many ways, from the identification of players with characters to the 3D environment that they explore. This "life-like" environment begets the need for a more realistic system of reputation. Because players' choices and actions are relatively unconstrained, especially when compared to a service like eBay, it is natural that reputation requires more subtle individual interpretation.

Though the systems could be improved, it appears that reputation supports WoW's purpose relatively well by creating trust between players. WoW is an unconditional success in terms of number of subscriptions and the amount of game-play users put in. The game has made it easy to broadcast reputation and show off one's achievements, which has proven a motivator to many players. Just as in an offline job, competition breeds many hours put towards advancement, and such obsession translates into significant profit for Blizzard. Reputation seems to operate fairly effectively among players as a signal of skill, supporting their purpose of advancing in the game, but it does not do a good job of broadcasting cooperative or defecting behavior. It is generally understood that players should be judged not only on their levels and items, but also on their speech and action. Players are also able to refer to the
social networks within the game, whether a guild or general chat channel, to check up on players, but this strategy does not always work. Those who play the game enough will be able to find out about other long-term players, which adds additional information to use in grouping decisions, but new players may be unable to do this. To augment these judgments, Blizzard might consider adding explicit measures by which players could rate the cooperation of those that they group with or label defectors in a way that ends up in a central repository like The Armory. This might also detract from the game and be open to some of the same problems that eBay faces.

Thus, reputation seems to work well-enough, most of the time, to do its job. Observational evidence suggests that many players can figure out whether another player is a Chinese gold farmer, an experienced warrior, or a newbie who has bought his character online. The costs associated with this evaluation, though they may be higher than if it were a perfect system, are not so high as to significantly harm gameplay. Nonetheless, information asymmetries do persist that have the potential to drive out "high-quality" players who do not consider reputation in their payoff structure. Better signaling mechanisms for cooperation would help link present and future payoffs to raise the cost of defection and link interactions together into a coordination super-game. Enlarging the coalition of players among which cooperative reputation is shared would also better support this goal. Still, the popularity of the game suggests that this problem is not so large as to drive away players, and that simple reputation mechanisms may do their job without adding unnecessary complexities.

As with eBay, WoW reputation is not portable beyond the fact that Blizzard publishes character information on The Armory. Moving to another MMORP in-
cludes a transaction cost of starting anew, and one's reputation on WoW cannot easily be transferred to a new gaming situation. This helps explain why players stay on the service for as long as they do. Inversely, players are unable to bring other reputations with them into the game. If they performed well in other games, bragging is their own signal. Though players may initiate bonds by discussing other games, it can be frustrating to have to start over. Similarly, merchants within the game cannot easily advertise their eBay reputation, which might increase trust when undertaking trades. WoW is a prime example of one useful part in the larger reputational puzzle, a piece of the pie that I will show in Chapter Five can be usefully amalgamated with others to paint a clearer picture of people's reputations in order to better support the signals that overcome information asymmetries.

The mixture of systems that makes up WoW's reputation mechanism makes it a rewarding test case. Players have many signals available to judge, from explicit level to characters' performances. Nonetheless, this required interpretation adds a cost to evaluating other players, and it is difficult to evaluate a potential playing partner's likelihood of cooperation. Complexity makes it harder to defraud the system, unlike on eBay, where one only has to deal with the feedback number to manufacture a reputation. Players still falsely signal in ways that increase the difficulty of making interaction decisions. However, as the theory predicted, players do protect their reputation because it is a reward to their egos and necessary to grouping with others to advance within the game. This makes them consider future as well as present payoffs, but the lack of credible punishments to defectors who can find other players on whom to prey makes trust signals less reliable than they could be. In short, reputation per-
forms admirably for broadcasting playing ability, but less so for broadcasting of co-operative behavior, making it an incomplete signal.
Mark Zuckerberg started Facebook in his Harvard dorm room in 2004. At first, membership on the site was restricted to those with a harvard.edu email address, but Facebook quickly began to let students with other college email addresses into their respective networks. High schools were then added, followed by companies, and in late 2006, the site opened up to anyone with a valid email address, no matter the domain. Facebook lets users create personalized pages—with information such as picture, name, educational history, email address, phone number, favorite movies, etc.—and then "friend" other users.* A user initiates a friend request to someone else on the site, who then has to confirm that the two "are, in fact, friends."154 Pairs of friends have the option of specifying the nature of their relationship, such as whether they work together, used to date, or met each other randomly. This information must be verified by both parties as well, and is then displayed for third parties to see.

Once two people are friends, they each have the ability to see the other's profile, though users may also choose to make their profile available to all Facebook users, users within their network (such as those with a wesleyan.edu email address), or friends-of-friends. The profile also includes listings of the user's friends and the number of friends the person viewing the profile page has in common with the user. In addition, there is a list of groups the user has joined on Facebook, ranging from serious ("Stop Genocide in Darfur") to silly ("Though Shalt Not Use Bruce Springs-

* A friend is Facebook's term for a connection, and I will use the terms interchangeably.
teen's Name In Vain’). Each profile page also has the option of a "wall," which is like a public message board for each user. On this wall, which is automatically en-abled and which many users choose to keep, connections can post short messages that everyone who can view the profile can see. Thus the profile becomes an e-represen-tation of its owner, and observational evidence suggests that people largely represent their identities accurately.

Figure 8: The author's Facebook profile page.
Facebook falls under the general category of a social networking site (SNS). SNSs have exploded in the past five years, though early versions began as early as 1997.\textsuperscript{155} Danah Boyd and Nicole Ellison offer a definition, echoed by others, of SNSs as "web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system."\textsuperscript{156} SNSs primarily enable users to make their social networks and web of relationship connections visible, both to themselves and others.\textsuperscript{157} This is in contrast to dating sites, which explicitly encourage strangers to meet, and other websites that do not focus on broadcasting users' true identities.

Facebook, the second most popular SNS in the world, is the runaway favorite of college students. Estimates of undergraduate use of Facebook in the United States range from 80%\textsuperscript{158} to 90%.\textsuperscript{159} Its popularity, especially among college undergraduates, is one of the aspects that makes Facebook so interesting to study. Many students spend hours a day on the site, checking up on friends' profiles, posting on walls, playing games, creating events, and sending private messages to other members. It also acts as many students' homepages, making it the first thing they see when they open a web browser. Such widespread adoption suggests the question, what is the purpose driving this use?

Facebook's purpose is multi-faceted, and as such, difficult to discern. On the one hand, entertainment is an objective, and many users go on the site to "hang-out" in much the same way they would offline. Social games are becoming popular on
Facebook, such as Scrabulous, a way to play Scrabble with one's friends. But even underlying this fun, there is the basic purpose of connection: there is no functional use of Facebook for a user who has not identified friends and accurately represented his or her social graph (the display of one's connections), so losing "friends" is a meaningful punishment. Facebook makes it easy to keep up with friends, present and old, as well as see with whom one may have a one- or two-degree connection. It is not uncommon to see postings on walls asking for information ("What was the name of that restaurant you mentioned last night?") or just merely saying hi ("Hey, haven't talked for a while. Sup?"). In this way, Facebook has a close connection to users' offline lives. Most people connect with those they already know, and Facebook's event system is used to plan parties and other functions. Anecdotal evidence suggests that using the site to network—connect with new people—is less popular, but users sometimes connect with friends-of-friends, especially if they know they have something in common, like attending the same school or working at the same company. Some may say that getting a large number of friends is the basic purpose, but connection strength must also be taken into account. Because most people friend others that they already know offline, deciding whether potential connections are who they seem to be, and not impostors, is critical. The company's purpose, then, is to support these activities by creating a website on which people feel safe connecting with others and broadcasting personal information, which leads to more users. The more eyeballs, so the business model goes, the more profit, and Facebook capitalizes on this by displaying advertisements.
Given these broad purposes, Facebook interaction is more difficult to model using game theory than eBay or WoW, especially because the meaning of reputation is less obvious than it is on these sites. Facebook is primarily meant to broadcast identity (the nouns that describe a person) rather than reputation (the adjectives that describe a person, such as trustworthiness and skill at a certain activity). Users can make reputational judgments (Is this person popular? Does he look like he would be reliable? Is she who she says she is?) about people on Facebook using the information on their profiles, but because this information is primarily used outside of the site, Facebook can hardly be said to have a built-in, designed reputation mechanism to track these kinds of judgments. Rather, reputation gleaned from Facebook is more like the word-of-mouth systems discussed in the theory section. But instead of relying on reputation being passed along by connections as it is in offline life, users rely primarily on the word of the person being evaluated, making Facebook's reputation system best described as "word-of-profile-page." Thus, there are two types of reputation on Facebook: that which is used to decide whether to connect with a person online (i.e. that suggests the person is who he or she seems to be), and that which is used in other forums (i.e. that suggests the person is trustworthy or fun, decisions that are useful in any relationship, online and off). To be effective, Facebook must support accurate identity information on the profile page that its users can use to make both kinds of reputation judgments.

This suggests the meaning of "cooperation" and "defection" in the context of Facebook. Cooperation is accurately representing one's identity so that others can make informed reputation judgments based on this information. Defection, on the
other hand, is falsely signaling identity information (how users do so will be dis-
cussed below) in order to get the benefits that come from connecting with others on
the site and "showing" them this information.* The benefits garnered from defection
of this kind may take two forms. First, connecting with others gives one access to
their personal information such as phone number, interests, and pictures. This infor-
mation could theoretically be used for harmful purposes. Second, because Facebook
users turn to the site to find out more about or check-up on people, lying about one's
group affiliations or educational and work history falsely signals identity information,
which others use to assess reputation (such as intelligence) and which may have im-
portant effects on users' offline lives. This creates a classic information asymmetry
that, if it were to get out of control, would lead to users distrusting the information
they glean from Facebook and ceasing to use the site, in an extrapolation of Akerlof's
theory discussed earlier. Luckily, the importance of Facebook connections builds in a
punishment for false signals, which will be examined in more detail below, making
users think hard before exaggerating too much on their profile pages or impersonating
someone else for fear of losing the connections that provide much of the site's pur-
pose. Thus, the present payoffs from lying on one's profile to make other people
think a certain way must be weighed against the future punishment that this false sig-
nal may lead to connections "de-friending" one. Therefore, Facebook activity (friend-
ing others and filling out one's profile) is best modeled by the coordination game dis-

* Other kinds of "defection" on the site include reposting the private information of others
elsewhere, defacing users' profile pages, or contacting others through the site for favors without
reciprocating. I will focus on the defection of misrepresenting oneself, however.
cussed in the theory section rather than a series of single shot prisoners' dilemma games.

Accurate identity information that can be used by others (such as to make reputational judgments) is needed to support Facebook's purpose discussed above. If the transactions cost of making an evaluation of someone's identity is too high, then the link will not happen. How users make this evaluation is explained below, but if false information too easily passes for truth, users will lose their trust in Facebook and stop using it. Similarly, if it is too costly to reliably signal identity, users will not populate the site. Cliff Lampe, Nicole Ellison, and Charles Steinfield note that "Just as in economic exchanges, search costs can easily influence the likelihood that a connection will be made." Without accurate identity, users won't be able to find one another and trust in others on the site will break down, driving away users and hindering the site's purposes. Luckily, Facebook (as an SNS) represents a well-functioning model of a reputation mechanism, in part because it mirrors how offline life actually works. People's identity information is judged and supported by the connections they have, just as someone might ask a friend to vouch for a new acquaintance in the professional world. This model of reputation has proven quite effective in supporting the trust required for the social connection purpose of Facebook, but the problem, as I will show, is that the information contained within it cannot yet be used effectively in other forums.
1. Facebook's Reputation System

Facebook has a complex reputation system that mirrors, in many ways, the way reputation is discerned and spread in everyday life. The public display of friend links and informational signaling on the profile interact to project users' identities. Facebook also has the feature of being able to check users' email address domains (the part after the @ sign) in order to let them into a particular network (a restricted group on the site). To do so, the site looks at the domain of a user's claimed email address and then confirms that the user actually owns this address by sending an email to it with a link that the user must click. For domains belonging to colleges and corporations that do not give out email addresses to anyone but those who belong to the organization, this can be an effective way of vetting one dimension of identity. But it is subordinate to and less complex than the profile and connections-based mechanism, to which I now turn.

Accepting a connection request implies that the user knows the requester or would like to know this person. Accepting a connection also implies that one believes that the person controlling the profile is the person represented on it. One way to make this judgment is by looking at to whom a potential connection is already connected. Connections have presumably viewed the profile and, by continuing to be Facebook friends with the person, implicitly verified that they believe the information on the profile to be true. Moreover, according to Tad Hogg and Lada Adamic, "each link in the network [acts] as an implicit recommendation for the person linked to." As punishment, they can sever the connection if the relationship goes sour or if they
see something on the profile that they do not want to be associated with (such as inappropriate pictures, stories of reckless drug use, or something they know to be untrue).\textsuperscript{162} There is also a sense of being watched that leads to a higher likelihood of users telling the truth on their profile.\textsuperscript{163} It is easy to get away with fibbing on a pseudonymous message board, but on a site like Facebook that requires real names, it may be embarrassing to have friends catch one in a blatant lie, thus raising the costs of lying (defection) above the benefits of potentially getting more friends. Lying can also cause one to lose already established friends, who, because it hurts them to be friends with someone who would betray their trust and lie, have the credible threat of ending their connection if they catch the falsehood.

Donath and Boyd note the use of connections for making "identity claims" reliable."\textsuperscript{164} One can find clues about a stranger based on their connections, especially if one is connected to these people as well. Of course, one is more likely to trust the opinions (as signaled by the link to another) of close friends, so if Facebook indicates that a potential friend has many connections in common, then this is a fairly reliable signal that his or her profile page accurately signals identity. Removing one's link to select others is a strategy to game the system by hiding the identity information that these connections signal, but then one loses the benefits of the network, which include getting job referrals and keeping in touch.\textsuperscript{165} Similarly, one could gain reputation by friending lots of people viewed as reliable, but these people are unlikely to accept connection requests, understanding the dangerous reputational effects this can
have in an SNS like Facebook, where connection is seen as recommendation. * Publicly displaying friends is thus an example of the handicap principle at work: in freely giving out this information, users are stating their confidence in the truth of their identities by wasting their right to the privacy of their connections and risking others' using this information to tell their friends about the things they do, good or bad.

While links may be the most reliable signal of reputation on Facebook, displayed profile elements are another, albeit conventional, signal similar to the performance-based reputation system in place in WoW. As discussed, users have a choice of how much information to display on their profile page, ranging from pictures and favorite quotes to phone number and physical address. When one user initiates a friend request to another, the requested user is able to view the requesting user's profile. This helps users evaluate potential connections' identities. There may also be reputational clues within the profile, such as school information or quotes illustrating a sense of humor. Fred Stutzman finds that name, academic classification, gender, email, friend network, and picture are the most likely fields to be included, while phone number is the least likely. 166 On average, he finds, Facebook users fill out 50% of the profile fields possible, suggesting that many users make an effort to provide identity information. Of course, some profile fields are more reliable: as discussed, joining a school or company network requires an email from that domain. Others, such as picture or interests, are much easier to fake. Pictures may seem like good sig-

* Interestingly, there is a fairly widespread phenomenon of people friending others, whom they don't know, with their same name. Nonetheless, the majority of connections are between people who know each other.
nals, though are obviously easy to fake, even by finding a picture of the person in
question in other users' Facebook albums. Nonetheless, the picture can be a better
signal when interpreted by the right people. Donath discusses the idea of "fashion" as
a signal, explaining that being up on certain trends signals a person's membership in a
particular group. This can be usefully extended to include profile information as
signaling "part of a certain person." If someone is trying to fake the profile of a per-
son and has even managed to trick others and therefore get plausible connections,
close offline friends may still be able to sniff out the fraud by looking at the picture or
other profile information. Even if the fraudster has done a good job, and the informa-
tion may have been true at some time, it may constitute an "old" identity. The picture
may be in a pose that close friends know the person would now never publicly dis-
play or include others in the photo with whom the person is no longer friends. Thus,
the only person who can truly fill out a profile, such that it "tricks" even the closest
friends, is generally the person in question. Of course, deceit is possible, a possibility
I expand upon in the problems section, but these commonly used profile signals lower
its incidence, especially when users are worried that lying will lead to lost
connections.

While not much of the information contained within the profile constitutes a
reliable signal, as a whole it offers a decent snapshot of users' identities, especially if
a potential connection knows what he or she is looking for and has the tools to inter-
pret the information. If a user is part of the right network, has a picture, and puts
down a few favorite books that one would associate with him or her, then there is a
good chance the person behind the profile is who they say they are. Of course, some
information signals more reliably; these profile elements include email address, network affiliation, and phone number. Phone numbers reliably signal, because a person checking up on a reputation can pick up the phone and dial to see if it is who they think, thus making phone numbers an example of the handicap principle. Email addresses accomplish a similar thing, but are less reliable if they are from a free email provider such as Gmail or Yahoo, because anyone can register under assumed names on these sites. Network affiliations, however, are the most reliable. Facebook evaluates users' right to access certain networks (such as school or company) based primarily on their email addresses. Alessandro Acquisti and Ralph Gross note that these signals increase "the expectations of validity of the personal information therein provided..." If users trust identity signals on Facebook, they will have more trust that users are who they say they are and thus be better equipped to use the information on profiles to make reputation judgments.

Limiting network membership adds value by enabling accurate signaling of identity information such as school attended, just as does the norm of not adding everyone under the sun as a friend. Donath suggests further increasing the amount of knowledge required to join networks, such as by asking what color the third floor lockers are for a high school network. Of course, even without a certain domain's email address, one can join Facebook and attempt to trick others in different ways using other profile fields. But empirical evidence supports the claim that more reliable profile elements (such as school) lead to better identity evaluation. According to Lampe, Ellison, and Steinfeld, following the idea that certain profile elements constitute reliable identity signals,
"Assuming that people will be more likely to articulate a relationship with...those who include these verifiable identity cues (such as high school) in their profile, we thus expect that the more these elements are included, the more "successful" one will be in the Facebook context, using size of friendship network as a proxy for success."\textsuperscript{172}

In this way, accurate identity signaling (which is needed to build trust on the site) becomes a reward in the form of more of the useful connections discussed above.

Those users who misrepresent themselves will be seen to have a bad reputation (as having a profile not corresponding to their identity) and be punished in the form of fewer friends. Thus, despite the fact that not all profile information is reliable, payoff analysis and evidence reveals that as a whole, the profile is an effective way to broadcast identity and increase trust.

Wasting this right to privacy on Facebook is another example of the handicap principle. By including private information on a page that, while not public, is likely accessible by many people (especially if the user has set their privacy setting to allow profile viewing by anyone in their network, as most do), users signal a desire to let others find them, so that they may build a network and garner all of its positive effects. Interestingly, Stutzman finds that users claim they are worried about privacy, but still let others see their profiles.\textsuperscript{173} Similarly, Acquisti and Gross note that users' privacy attitudes suggest they would not use the site as much as they do. The authors find that users worry more about privacy than terrorism or global warning, yet still take few steps to make their information on Facebook accessible only by friends.\textsuperscript{174} Why the disconnect? A satisfactory answer would take more empirical research, but it seems that users see the benefits of an active and accurate Facebook membership as
so great that they are willing to look past their privacy concerns. Making their profiles more private is one option (one that this author takes) but by doing so, users lose out on many of the signaling options discussed above and raise search costs for others, thus limiting their ability to make more valuable connections. In line with Lampe, Elison, and Steinfeld's research on profile information leading to more friends, users make a decision about how much privacy to give up in order to signal to others and usefully increase the size of their network. Facebook thus incentivizes its users to broadcast accurate and public signals to build trust on, and therefore use of, the site.

In short, people use profile information to judge others' identities on Facebook. A quick glance at name, picture, and common friends can yield a relatively good snapshot of a person, from which one can make a judgment as to whether or not to trust that person's identity enough to add him or her as a connection. Consequently, connections and profile information combine to produce identity information that can be used to judge reputation, though this requires more interpretation on the part of the viewer and therefore has higher use costs than an eBay feedback score. By encouraging accurate identity signals, however, Facebook lowers the risk aversion of users when adding new connections, thus supporting the site's purpose of providing an accurate representation of users' social graphs. Signals also link present to future payoffs to increase cooperation in the form of truthful identity broadcasting.
2. Problems

Facebook's reputation system, though more accurate in projecting users' identities than the other case studies we have encountered, nonetheless has problems of its own. These problems fall into the categories of identity fraud and how users can make use of reputation information gleaned from the site. Overall, Facebook does a relatively good job of enabling accurate identity determination, but how it enables users to employ this information for purposes besides making connections and in contexts outside of the site is more problematic.

Identity fraud is bound to be an issue on any social network service, but the protections discussed above, as well as the norms and rules of Facebook, have largely stemmed this problem. The Facebook terms of use state that users may not:

- register for more than one User account, register for a User account on behalf of an individual other than yourself, or register for a User account on behalf of any group or entity; [or] impersonate any person or entity, or falsely state or otherwise misrepresent yourself, your age or your affiliation with any person or entity.\(^{175}\)

Recently, the staff of the Wesleyan Argus school newspaper, who wanted a Facebook page on which to post photos from the newspaper and other information, had their account suspended for not being an individual person, but managed to make a new page under the pseudonym A.R. Gus.\(^{176}\) Of course, Facebook cannot actively monitor every profile on the site, but there are many reported instances of people being kicked off the service for using a fake name or impersonating a well-known celebrity.\(^{177}\) This is in contrast to other SNSs, such as MySpace or Friendster, that have many profile pages representing bands, celebrities, and fictional entities.
High profile frauds do slip through, however. In late 2007, a fake profile claiming to be Bilawal Bhutto was discovered on Facebook. A prankster decided to make the phony page, which, given 19 year-old Bhutto's new position as chairman of his late mother's Pakistani political party, had dangerous potential. Bhutto and others discovered the fraud, though it fooled major newspapers for a while, and contacted Facebook to have the page taken down, a request with which it complied. A statement Facebook released about the incident reads like a grocery list of reputation mechanisms in effect on the site:

After investigating the accounts in question, we have disabled both Facebook profiles associated with names Bilawal Bhutto and Bilawal Bhutto Zardari. We found they were not authentic and violated the site’s Terms of Use. Typically we examine a range of criteria to determine whether a profile is authentic, including reports from users, profile content, the e-mail associated with an account, length of time the account has been open and network affiliations.178

In Morocco, Fouad Mourtada was arrested for posing as the King’s brother on Facebook, a crime which resulted in three years in prison and for which he was later pardoned.179 Notwithstanding any moral judgment of this court decision, it shows, along with the Bhutto incident, that people take Facebook fraud seriously, especially because profiles are assumed to be more accurate than on other sites.

Despite these high profile cases, evidence suggests that Facebook's reputation system has done a good job of enabling trust on the site. Again, "code is law." Facebook's reputation system, through its system of public links and profile fields, encourages people to be truthful. The underlying code of the site allows false identity, but the surrounding reputation mechanism, also coded into the site, stems it.
course, the norms of the site also uphold truthful identity, perhaps because it was first populated by college students intent on having useful access to connections, but these would not be able to operate without the underlying reputation system that enables reliable identity signals. It is easy to register with a fake name, but then one does not get most of the benefits from the site. In this way, Facebook has done an exceptional job of linking truthfulness to reward in its reputation mechanism, and research shows that it is working.

Acquisti and Gross find that most people provide accurate information on their profiles. They find that users trust their friends on Facebook to a high degree and even trust others in their network, even if they are not known, at a fairly high level. They also find that users trust those not connected to them on Facebook at a much higher rate than unconnected users on other SNSs such as MySpace or Friendster, if they also use these services. These figures suggest that Facebook is doing an effective job of encouraging trust among its users, which has led to a high rate of adoption of the service and a fulfilled purpose of displaying accurate social graphs and enabling users to signal their identity for others to make reputation judgments with.

Another, perhaps more fundamental problem with Facebook is the relationship it creates, or fails to create, between identity and reputation. Recall, from earlier, that identity is a claim about oneself, including such information as name (or pseudo-

--

* Users, on average, rated their trust of friends on Facebook at 5.62/7, those within their network at 4.35/7, unconnected users at 3.29/7, unconnected users on MySpace at 2.72/7, and unconnected users on Friendster at 2.32/7 (Acquisti and Gross, p. 18).
nym), privileges within a group, and history of action. Sites use identity to decide what actions users can perform and what information they can view, while users employ available identity information about others to make decisions about who they are. Reputation, on the other hand, constitutes a story about a person, often constructed using past actions and identity information, and makes claims about meta-identity, such as whether a person is trustworthy and likely to cooperate or behave with another's best interests in mind. Assuming that offline friendship is a sign that one trusts the other person, a Facebook connection may be seen as stating that a person is trustworthy. To some extent this implicit recommendation may be the case, but bonds can be so weak among Facebook friends that making such extrapolations is meaningless. Thus, a Facebook profile almost entirely represents identity, rather than reputation, information. It is up to other users to extract reputational claims from this data, such as from profile information that suggests trustworthiness or from the user's connections to other trusted individuals.

Because it is mostly used to promulgate honest identity, Facebook reputation thus has important effects on outside life. This paper does not claim to "make the jump" between online and offline in any meaningful way, so I will not delve deeply into this connection, but it is a vital one to mention. Job interviewers are beginning to use Facebook to check up on candidates, as are people about to go on a date. A Facebook profile can often be a good, if shallow, representation of a person. Even though some choose to fill out the profile fields ironically, this too can tell a lot about a user, especially if the page viewer understands the humor being used. Though reputational claims of trustworthiness may not have widespread uses on the site itself,
they do have important applications for users' other activities, both online and offline, making most users take quite seriously the information they display on their profile and the people they befriend on the site. At the same time, because reputation (such as whether a person seems trustworthy or funny) gleaned from Facebook is also used outside of the context of the site, there is the danger that punishments that happen outside of the scope of Facebook are not reflected back onto the site. If someone makes a reputational judgment using Facebook identity information that turns out to be wrong, there is no simple way to "tell" others. The betrayed person may remember to cancel their Facebook friendship with the untrustworthy person, but other users of the site may not notice this. In other words, the reward from proving identity through a connection may be greater than the punishment from losing that connection. Facebook's intimate connection to users' offline lives (because real names are almost always used, profile information, connections, and any reputation judgments made using this information are associated with users offline) thus makes it a force worthy of careful consideration, but one that is sometimes difficult to separate from this offline activity.

Because Facebook presents such a good snapshot of a person's identity and connections, from which others can make reputational judgments, it would be useful if its information could be more easily used in other contexts. Information on Facebook, such as contact lists, cannot be automatically exported from the site to other SNSs or websites. This will be the subject of the next chapter, which discusses ways to consolidate online reputation in order to more fully support widespread trust, but suffice it to say here that Facebook information is markedly not portable. This is
noteworthy, because Facebook identity information would often be useful in informing identity and reputation in other contexts in order to enlarge the "coalition" of these contexts to include reliable information from offline life.

Making reputation easier to broadcast would make the information that Facebook houses about its users more useful in other contexts. One way of doing so would be to always display how many, and which, contacts one has in common with another user, even if one is not friends with that user. Reputation could also be made a more explicit feature of Facebook, though it is unclear whether this would disturb the social mores on the site. Another SNS, LinkedIn.com, which focuses on professional networking, gives users the option to "recommend" others for work they did together. Given the more general focus of Facebook, this may not be the best solution to making reputational information more accessible, but third-party applications, such as "Reputation" and "Rate me!" let users rate one another on a variety of metrics, including trustworthiness. How widely and in what way such information is used is unclear and would require further research. Of course, the information would still only accessible by those who have access to a person's profile, often only after they have become friends and therefore established basic trust. Users might find it useful to be able to make this information publicly visible, such as next to one's name and picture in Facebook search results. More detailed positions on how to do this will be discussed in the next chapter, but notice how encompassing, yet unusable, Facebook identity and reputation information can be.
3. Conclusions

Facebook is the foremost example of implicit reputation on the Internet. Profile elements represent conventional signals, though revealing verifiable information such as phone number, network, and email address borders on reliable by requiring wasting the right to privacy as a signal of truthful identity. Publicly displaying one's connections is a similar example of the handicap principle, by giving others access to one's social network and therefore the ability to check up on one. Using connections for a judgment of identity is most reliable, because it is harder to trick many users (especially those close to the person whose identity is stolen, and therefore represent a stronger claim of veracity) into believing the identity theft. Connections thus contribute most to one's reputation as displaying truthful identity information on Facebook.

This system of reputation is remarkably similar to how things work offline. Strangers meeting at a party attempt to establish a common ground, and when they find people they both know, it is easier to make a positive judgment of the other person and build trust. As the old saying goes, "a friend of yours is a friend of mine." Thus, Facebook connections might be seen as an implicit recommendation of a person and their trustworthiness, and therefore act as explicit reputational information. This understanding that a connection implies a recommendation, and the benefits that people on Facebook get by being connected to one another, makes severing Facebook friendship a credible punishment. Knowing that connections will see the identity information they put on their profile page and others to whom they are connected, users
will thus be more likely to maintain an honest reputation by telling the truth. For the most part, however, users are left to make their own reputational judgments about others on the site based on their identity attributes, which include profile elements and connections, especially those held in common.

This required reputational calculus makes information decidedly difficult to port to other contexts without the leg-work of evaluating users on Facebook and then manually bringing this judgment elsewhere. How this translation would even be done is unclear. Part of the effectiveness of Facebook, particularly in its resistance to fraud, comes from its complexity and the requirement that users interpret the identity information they find to make their own reputation stories about others. As Bob Blakley has noted, not everyone uses even simple rating information in the same way; this effect is heightened when reputation is as complex and interpretable as it is on Facebook. Because a person can have multiple reputations (as a good student, a reliable friend, etc.), and reputation depends on the perspective of the interpreter, this leads to a better, though more costly, judgment than the simple number ratings in place on sites like eBay. Facebook reputation would be impossible to boil down to a single number: number of friends does not suffice, nor does number of friends in common, though both are useful numbers to examine in deciding whether to trust a signaled identity. True reputation metrics, such as trustworthiness, can only be discerned once one has other interactions with the person in question, whether offline or on Facebook through such channels as messages or wall posts. There is no obvious way to quantify these.
Even so, information gleaned from Facebook may be useful when combined with other data. To give a preview of the next chapter, imagine the usefulness of being able to see how many friends a potential eBay seller has in common with one, or their network, and then being given a link to their profile page to make further judgments. This portability would help increase the size of the "coalition" and the amount of identity information available to reputation requesters in such a way as to potentially lower their risk aversion. Though Facebook reputation may not be able to stand on its own in many situations, due to the difficulty of interpreting it and porting its data, it acts as an effective support for and context in which to interpret other reputation information. It can link online identity to offline and other contexts that the reputation-requester may have more experience with, such as a college network or a Fortune 500 company. As such, Facebook offers a useful framework for other identity information to operate within, though as of now, this cannot happen because Facebook data is not accessible to other SNSs or other classes of sites. I will examine in the next chapter how this might be done.

*****

The above case studies have empirically supported the predictions that the theory made about the roles of identity and reputation in supporting the trust that is necessary to support online social interaction. eBay uses explicit measures to accomplish this, WoW relies on character level and performance, and Facebook utilizes profile information and users' stated connections to others. Overall, these three test cases
give practical examples of how changing players' payoff structures by introducing reputation can link present to future games. This shift replaces the likelihood of defection in a prisoners' dilemma with the possibility of cooperation in a coordination super-game. The trust that reputation signals furthermore lowers risk aversion in such a way that players feel comfortable cooperating (in whatever way is meaningful in the current context). Reputation's power is so strong that it can overcome the lack of physical cues online to support users' trusting one another when trading with real money, grouping in a game that takes up large amounts of their time, and believing that a profile page belongs to whom it seems.

Reputation, as we have seen, is imperfect. All three systems that have been examined are open to fraud in such a way that threatens to break-down trust. So far this has not happened on a large enough scale to render the systems useless, but the problems section in each chapter above illustrates that this is a possibility. The theory would suggest that one strategy to increase trust is to further reduce risk aversion by increasing the reliability and number of signals available to players. The best way to do so, I believe, is to link reputational contexts in such a way that increases the size of Greif's "coalition" and more strongly forces players to consider both present and future payoffs. It is to this possibility that I now turn.
In the previous three cases studies, which cover the three main reputation models that operate among between individuals on the Internet, data is not portable as a reliable signal. As Kim Cameron explains, because "a way to know who and what you are connecting to...is missing [on the Internet], everyone offering an Internet service has had to come up with a workaround. It is fair to say that today's Internet, absent an identity layer, is based on a patchwork of identity one-offs. \textsuperscript{186} In other words, reputation and identity information gleaned from eBay, WoW, and Facebook cannot be exported to other websites in such a way that the information can be proven to be dependable. If I am attempting to establish trust on a message board, I might point someone towards my eBay seller page, WoW character profile on The Armory, or Facebook page to prove some facts about my identity and reputation, but it is difficult to prove that the pages "belong" to me. There are ways to do so, but they are costly, and as such unlikely to be used.

One way is to pass a "token" between parties in all forums that require linking. Because this capability is not built into the Internet or Web, it is a costly endeavor. An example will suffice to illustrate its complexity. Imagine that two people meet on a gaming message board and want to prove to each other that they have level 70 characters in WoW, because they like to brag or are having an argument about
playing strategy, and that they are respected sellers on eBay, because they would like
to sell each other rare items but want to make sure the other is trustworthy. First, the
two would have to exchange emails. Then, over this communication channel, they
would have to decide on some sort of password and response to use as the identity to-
kens on the other sites. Then, they would have to find each other in WoW (which
would mean being on the same playing server, of which there are dozens, and if they
are not, transferring their characters, which can incur a cost) and exchange these to-
kens via private chat. It would then be up to each to judge the other's character in the
ways discussed in the WoW case study, some of which are unreliable. If each was
satisfied with their interpretation of the other's WoW reputation, then this would suf-
fice to export the trust garnered from this context. On eBay, the two would have to
undertake a similar process. If the parties took these steps, then they would have
linked the pseudonym used on the message board to their email addresses, WoW
characters, and eBay seller profiles. The process of doing so, suffice it to say, is quite
costly, and only serves to link the pseudonyms, not the person behind them, who
could be a friend pretending to be either party.

Even if the person behind all of the identities is the same, and the parties are
able to construct a combined reputation of one another from all the desired contexts,
this does not imply any guarantee of future action, especially if the parties interact
outside of the linked reputation arenas. For example, if the two decide to trade items
in WoW based on each other's impressive seller ratings on eBay, there is no credible
punishment if the trade goes sour. The defector's eBay score is not affected, and the
only recourse the defrauded cooperator has is to complain about the other in WoW
and on the message board, but the likelihood of this making a difference is small, given the scale of WoW and the ease with which the defector could disappear from the message board. Thus, the process of transporting reputations is time-consuming and the lack of credible punishments means that future games are not linked to present ones, making defection in the resulting prisoners’ dilemma a rational choice.

The failures of these sites to provide acceptable reputation mechanisms that enable low cost trust, support the purpose of each site, and are complex enough to be resistant to fraud suggests the need for a consolidated system that creates the possibility of reputation across the Internet. Such a consolidated system would provide the tools that anyone could use to build up trust with others in any social interaction, such as playing together in a game, introducing one another to other trusted contacts, trading, or other actions that require reputation to encourage cooperation, curb free-riding, and prevent defection through the use of credible punishments in the form of injured reputation. OASIS, the international open standards group, announced in April 2008 that they are forming the Open Reputation Management Systems Technical Committee to define the standards for what shape these tools will take. The following excerpt from their April press release echoes my assertions of the problems that a consolidated system would have to overcome:

Because the majority of existing on-line rating, scoring and reputation mechanisms have been developed by private companies using proprietary schemas, there is currently no common method to query, store, aggregate, or verify claims between systems. The different sources of reputation data — user feedback channels (product ratings, comment forms), online user profiles, etc. — are each uniquely susceptible to bad actors, manipulation of data for specific purposes, and spammers.187
James Bryce Clark, a member of the committee, explains that the system "will pro-
vide standard ways to express assertions, evaluations and comparisons of rating and
reputation data, making it easier to consume that data [and] aggregate it." How the
group will define the technical standards is unclear, but the fact that a committee has
been formed to address issues of reputation consolidation illustrates how important
industry considers solving these problems of siloed identity.

A consolidated reputation system can usefully be thought of as an automated
version of the identity token passing discussed above. Rather than making users jump
through the hoops of creating their own shared secrets and exchanging them in the
contexts they require, this mechanism would automatically pass the tokens that prove
linked identity, thus making it harder to game. This proposed mechanism is also
more useful than the identity token, because it enables credible punishments by "read-
write" ability. For example, in the example above, a consolidated system might auto-
matically pull in eBay information about a character wishing to make an item trade in
WoW, and then enable the buyer to rate the seller within WoW such that it is reflect-
ed in his or her eBay seller profile. The identity linking possibilities are theoretically
endless, and more examples will be discussed later in this chapter. In short, the at-
tractiveness of such a system is the ability to consolidate formerly siloed identity and
reputation information (i.e. that which is trapped and only able to be used within the
context of one site) in a way that strengthens the trust-building capability of users' re-
putations while allowing services to preserve whatever "flavor" (explicit, implicit, or
performance-based) that they decide best supports their purpose. Users, however, do
not always want to link their identities, and drawbacks of the system include prevent-
ing the Internet from continuing to support multiple identities and pseudonymous interaction.

This service-specific choice brings up an important qualification: I am not proposing to replace all reputation systems on the Internet with one consolidated mechanism, nor am I arguing that all of one's online activities must be linked. Doing so has advantages and drawbacks. I firmly believe that the ability to interact anonymously (and pseudonymously) should be preserved. Service providers and users should have a choice of what level of consolidation and identity verification they require in different social interactions as well as how much anonymity to preserve, a critical privacy choice that will be discussed further. Privacy and the possibility of anonymity and unlinked contexts should be maintained; in other words, just because there is a consolidated system that allows users to broadcast and share identity information between contexts does not mean it always has to be used, but for many sites, the system's benefits will be difficult to pass up. As will be seen, a consolidated system promises to reduce transaction costs and increase the number and reliability of signals users can broadcast and utilize. Its negatives include a reduction in privacy and the decreased ability to maintain multiple identities and reputations according to one's needs or fancy. I believe that the benefits outweigh the downsides in certain situations that I will discuss further below, but the choice of how to link their identities must be left up to the users, who should be able to decide whether or not they want to use a consolidated system, just as users have the choice of whether or not to reveal their phone number on Facebook.
Consolidating the identity and reputation systems of different sites is a step that the theory predicts should increase the prospects of trust in social interactions. Recall the information asymmetry problem that Akerlof posits in "The Market for Lemons." When people do not have enough information about those with whom they interact, deception results and leads to distrust and fewer interactions results; that is, low quality goods (or, in social interactions, defectors) drive out high quality goods (or, trustworthy individuals). Information asymmetry is one of the central problems in the prisoners' dilemma that leads to the defection of risk averse agents who are afraid that their opponents will also defect. As Spence and evolutionary biologists show, signaling is one way to overcome this problem and broadcast information about oneself to other parties. By building trust that the other will cooperate, signals offer the possibility of cooperation. One such signal is reputation, which has the property of being able to affect the future payoffs of the person to whom it is attached. Those who defect in the present period will find it harder to build trust to support interaction in the future, thereby enabling reputation to link present to future payoffs in order to create the coordination super-game discussed in the theory. One way of multiplying reputation's power, Greif suggested, is to expand the size of the group of people (the coalition) in which it matters. Consolidating identity and reputation systems potentially fulfills this task and therefore enables reliable signals and robust trust.

In this chapter, I will discuss a few of the emerging standards that propose to solve some of the problems identified here, particularly how to link dispersed online identities and make measures of one's reputation portable. I will then describe the
need for extending the "identity layer" that many\textsuperscript{189} have proposed adding to the Internet to include a "reputation layer" that makes interpretation of people's meta-characteristics cheaper and therefore better supports trust in short-term social interactions that might not otherwise have the luxury of being supported by reputation. I call this novel system the \textit{distributed oracle}, because it enables the querying and use of identity and reputation information between contexts but does not require this data to be stored in a central location. As such, it consolidates access to this information, while allowing it to exist in the distributed contexts that it does now. I will conclude by discussing the problems with linking identities to form a consolidated reputation system, most of which can be solved by following Kim Cameron's first law of identity, which requires "user control and consent."\textsuperscript{190} This ensures that websites are able to keep their individual "flavors," which sometimes require anonymity, and that users are able to maintain multiple identities, a capability that should still be preserved on the Internet.

\section{Emerging Consolidation Standards and Their Problems}

Many of the problems I described above have been noticed by Internet industry players. Companies such as Microsoft and non-profit groups have been attempting to define and implement what is termed an "identity layer" for the Internet.\textsuperscript{191} An
identity layer is simply understood as being like the World Wide Web, a set of protocols built over and using the protocols of the Internet. Once such protocols are defined, any website can use them to share identity and reputation information with other services without having to negotiate the specific terms and form of the exchange each time. These protocols are commonly called a "layer," because they are used exclusively to pass a certain kind of information, in this case identity and reputation information, and can operate independently of other layers, such as the World Wide Web itself.

This architecture of connection is similar to the gas tank of a car. As long as the hole through which gas is pumped is standardized, people will be able to get gas for their car at any station. The gas pump does not care how fuel gets from the car's gas tank to the engine, how much fuel the tank can hold, or what kind of car it is. It just needs to be able to fit into the hole in the side of the car where gas is deposited. Similarly, the car does not care who owns the gas pump or where the gas came from. As long as the right kind of gas comes through the tube, it is happy. Standardizing the pieces needed for intake ensures that the provider (pump) and receiver (car) are agnostic towards each other.**

An identity layer would accomplish a similar purpose on the Internet. No matter what the underlying structure of either party is (a website or person), as long as

* Note that I use the standard term "identity layer" in order to keep my vocabulary consistent with what academics and professionals use, but this set of protocols would also be able to carry reputation information, as I will discuss in the next section. That is, my identity layer is more general than the term as it is used by the technology industry, and has uses beyond authentication and logging into websites.

** This is similar to what economists refer to as "homogeneity of goods."
both can hook themselves up to the identity layer, they are able to exchange identity and reputation information. In other words, the identity layer is not the form that users fill out on Facebook to fill out their profile, but rather the pipe that allows porting this information to MySpace. How this power is used is up to individual sites and is discussed further below. Of course, just because it is there does not mean that it has to be used, just as the owner of the car does not have to buy 93 octane gas if it is not the correct fuel for her car or might not trust the gas sold out of someone's backyard. Each party should be able to assess the quality of the information they are receiving and with whom they are sharing their own information. This ability is fundamental to preserving choice while "owning" one's reputation online, because when it is siloed within individual sites (as it is now) one cannot leave these sites or begin using new ones without incurring a cost. As we have seen, losing this ability also has important implications for the capability of sites to uphold the trust needed to support their purposes.

In this section, I examine two standards, openID and Social Graph, that make significant first steps in enabling this agnostic identity layer. Note that both are "open" standards, meaning that anyone can use them without paying and can look at their underlying source code. Clippinger notes that "no government or international body or single company can do this [design an identity layer] on their own," and past experience (such as in Microsoft's failed Passport system) has proven him correct. The openness of openID and Social Graph, as well as the roles they play, make them good candidates from which to bootstrap an identity layer, but they too have their problems. Note that eBay, Facebook, and WoW are not openID or Social Graph en-
abled (for the time being), so examples of how they are used will come from other websites.

openID is the closest protocol to a consolidated reputation system on the Internet. An openID is a URL (universal resource locator, like the www.google.com address that one types into the location bar of a web browser) that is used to identity its owner across any and all openID enabled sites. Looking at the process of getting and using an openID will make this clearer. First, one either signs up for an openID from a dedicated provider such as myOpenID.com, makes one's personal website an openID, or uses an openID that one may already have, which is the case for everyone who has an AOL or Yahoo username. Using this openID, one can then log into any openID enabled site without having to sign-up again. When using openID on a website, the site passes the agent logging in to a "handshake" site maintained by the agent's original openID provider at which to provide the username and password used on this original site.* The identity provider then passess a message to the new, requesting site along the lines of "this person does indeed own the domain they just typed in, so you should let them into your site with these credentials."

The real power of openID, however, comes in its role as an identity linker rather than as just a way to consolidate one's usernames and passwords into one openID. According to the openID foundation, "OpenID...realizes that people are already creating identities for themselves whether it be at their blog, photostream, profile page, etc."194 In a similar process to the messy token passing I described in the in-

* Cameron sees this as a problem because all of one's identities accessed by openID are only protected by a single username/password and thus more vulnerable to hacking.
roduction to this chapter, openID, according to Cameron, "gives us common identifiers for public personas that we can use across multiple websites—and a way to prove that we really own them," rather than "the cacophony of 'screen-names' we have today - screen-names [that are] in bondage, prisoners of each site" such as eBay, WoW, and Facebook. This is of significant interest as the solution to the problem of linking multiple identities, each with its own reputation, online. Using an openID to login and comment on blogs and as one's identifier on a host of websites provides a reliable signal that the same person is behind all of those presences. Furthermore, one can associate one's openID with any website that has editable HTML by adding two lines of code. This provides a reliable signal that the page is controlled by the same person who controls the openID, using a system of reciprocal linking that plays the role of passing a token.

An example will illustrate how this works. I have enabled my personal website, www.alexrosen.com, to act as my openID. So when I post a comment on a blog hosted by Blogger or Wordpress using my openID, anyone reading this comment can be sure it came from the person who owns www.alexrosen.com.* On this site, I could have confirmed links to other pages I can prove ownership of with openID, such as my Flickr photo account. These are reliable identity signals, but there can also be conventional signals, such as links to my Facebook page, which does not support

* Of course, this is not necessarily Alex Rosen, because anyone could register this domain, but friends of mine could be pretty sure that it was me by reading the content on the website. Strangers could also get a better idea of my interests and biases by reading www.alexrosen.com to put my post in context. In addition, "owns" does not necessarily imply legal ownership, but rather "use of." For example, I may not "own" the domain alexrosen.claimid.com, but the site operators have given me rights to use it as my identity page.
openID. There are also services, such as claimID.net and myOpenID.com, that provide this service automatically by giving one an openID that corresponds to a page on their server with identity information and confirmed links. Thus, openID provides an easier (though not perfect, because it only works when one has access to the HTML code on non-openID enabled websites) way to link online identities, which provides more identity information to others with which to make reputational decisions.

openID is a promising emerging standard that helps link identities across the Internet. It does not, however, fully solve the problem of enabling an identity and reputation layer online. First, it only links identity information and leaves reputation judgments up to the requester. Second, it does not have any access controls: one's openID is either public or does not exist. Finally, it can only read identity information, but update it, between contexts. These are problems that I will show can be solved by the more powerful distributed oracle discussed later in this chapter.

The second emerging standard that proposes to link identities and make reputation consolidation easier is Google's new Social Graph API (application programming interface), released in February 2008. Social Graph, which takes its name from the phrase used to describe the web of connections inherent in social networks, proposes to make this information portable across SNSs. To do so, Social Graph provides a simple way to search and gather the statements of these relationships that are already tagged on much of the Internet using the Friend-of-a-Friend (FOAF) and Extensible Friends Network (XFN) formats. These are open protocols by which users (or websites) can manually or automatically tag links with relationship meta-data
such as "me" or "friend." Social Graph automatically indexes and analyzes these tags that broadcast a connection to create an Internet-wide social graph of a person that can then be pulled into any context, such as an SNS. The "me" tag is special and must be reciprocated, so that for someone to assert that two sites both belong to them, each site must have a "me" link to the other. This is similar to the openID framework that enables these types of identity ownership assertions, but without the need for a "home-base" URL and with searching capability built in.

An example will help illustrate the power of the Social Graph API. Imagine that Dick is friends with Jane on Twitter, an SNS focused on sharing short "status" messages about what its users are doing. Dick decides he would like to join LiveJournal, another SNS that is more like a blog, as RichardII (because the username Dick was taken), but does not find the site useful without connecting to friends on it. He is also unsure of how to evaluate the reputation of potential connections on LiveJournal to create the trust that he would require to start building these connections.

What is Dick to do? Enter Social Graph. It turns out that Jane is already on LiveJournal, but her username is JaneEyre, so Dick may not have ever found her. By analyzing the tags on Twitter and LiveJournal, Social Graph knows that Twitter Jane is the same person as LiveJournal JaneEyre and that since Dick is friends with Twitter Jane, he may want to connect with LiveJournal JaneEyre. LiveJournal uses Social

* "Friend" can be more specific, because XFN includes tags such as contact, acquaintance, friend, met, co-worker, colleague, co-resident, neighbor, child, parent, sibling, spouse, kin, muse, crush, date, and sweetheart. Others could easily be added to better describe relationships (Google, 2008).

** Social Graph can even read openID information to ascertain "me" links.
Graph to give Dick this information, who can then connect with Jane on Live Journal, if he so chooses. Consider the following chart illustrating these principles:

![Diagram showing connections between Twitter and LiveJournal]

Figure 9: The connections of Dick and Jane across sites.

RichardII and JaneEyre are then able to quickly evaluate each other based on their respective Twitter reputations and instantly port the trust from this context to LiveJournal, which reduces evaluation costs for each of them and enlarges the coalition from which they can discern reputation data.

Social Graph makes it both simpler to evaluate others' reputations by enabling easy visualization of their connections and identities that are dispersed over the Inter-
net, and also enables websites to support their purposes by reaching a critical mass of users without the costs to users being as high as if they had to start from scratch. Google recognizes both of these benefits, but it is still unclear what shape the former will take. The company has released simple testing tools to help one visualize social graphs, but the real value will likely come as third-party companies develop more interesting applications for the data, such as finding all of a user's Internet-wide connections for use in making reputation judgments. At this time, less than three months after Social Graph's release, it is simply too early to tell, but based on Greif's theory, the opportunity presented by linking implicit reputations and scattered identities is enormous. By lowering costs of evaluating trust and finding connections, Social Graph will also likely give a boost to niche SNSs that offer useful services but have trouble building up a user base, such as Dopplr.com, which shares friends' itineraries so they can meet up with one another or friends-of-friends when traveling. Not having to evaluate and find new contacts would make this a much more attractive service.

It is important to highlight some of the privacy issues at stake, and I will return to this topic later in this chapter. Although Social Graph has significant potential in lowering the cost of trust by making identity and reputation easier to ascertain, there is the argument that this information violates privacy and opens a Pandora's box of relationships that people would prefer to keep private. Some argue that a person's social graph requires the context of the site it resides on and should not be shared among contexts. Nonetheless, it should be noted that Social Graph only indexes

*As Brad Fitzpatrick notes, "People are getting sick of registering and re-declaring their friends [and I would add, other information] on every site" (Fitzpatrick, 2007).
public data and as such cannot get into walled gardens such as Facebook. Tim O'Reilly submits the importance of the tool in "rais[ing] people's awareness, and ultimately their behavior." Many do not realize that this reputational data is freely available on the Internet, so by making it more accessible, Social Graph makes it usable by more people. It also encourages them to think carefully about the connections they are making, and makes the signaling power of connections stronger by invoking the handicap principle of wasting the ability to keep one's friends private. Thus, while privacy is critical to examine in relation to Social Graph, I believe the tool's benefits outweigh its costs.

Other systems do exist.* For example, Nishant Kaushik suggests using the Facebook platform as a springboard for consolidating reputation, given its robust implicit reputation system, norm of truthful identity, and the fact that many already turn to it (by manually going to the site and looking up users) for identity and reputation information. For certain situations, I believe that this is a good solution, and Facebook (as well as other SNSs) should be an optional part of any consolidation; however, on its own Facebook does not solve the present problem for two reasons. First, it is a closed, proprietary, and commercial system. This means that Facebook would own one's identity and reputation information even when it was used on other sites, which puts dangerous power into the company's hands. Second, the norm of truthful identity, while useful in some situations, gives away too much information in others,

* See also other identity management frameworks such as Higgins, an open source identity framework spearheaded by the Eclipse Foundation (Eclipse, 2008), Microsoft's CardSpace (Microsoft, 2007), and the forthcoming OASIS Open Reputation Management System (Cover, 2008).
and making Facebook the foundation of the system would prevent pseudonymous interaction. There are many situations where anonymous identity is useful, or only generalized identification is needed (such as this person lives is over 21, rather than this person is 27). When this type of meta-information is appropriate and why it is useful will be discussed further below. It may be possible to gerry-rig Facebook to give out this type of generalized identification, but keeping the data only in Facebook raises problems of privacy and centralization.

openID and Social Graph are flexible standards with the potential to link identities across formerly siloed websites, thus making it easier to evaluate reputation and create trust between interacting parties. Neither of them, however, goes far enough. Even when their use is combined, they do not solve the central problem of this thesis: how to use reputation information to support trust across websites in such a way that is resistant to fraud and efficient enough to be used by actors for temporary relationships that do not succumb to the pitfalls of the prisoners' dilemma. The two reveal three main problems that a consolidated system (such as my distributed oracle) would have to overcome: First, both are "read-only" one-way streets and do not link identities strongly enough. Second, and related, they do not do a good job broadcasting reputation, focusing instead on identity.* Third, the controls they offer are not fine-grained and do not easily support generalized information or uses in which a per-

* In other words, it is up to the requester to make reputational judgments using the identity information gleaned from either. While WoW and Facebook have shown that this act of interpretation is a critical part of reputation, it might be useful to have a more explicit measure. One could imagine a PeopleRank attached to Social Graph that estimated a "reputation score" based on the number of links a person has and to whom they are linked. Whether such a measure would be useful or accurate (or even what accuracy would mean in this context) would require significant research and empirical data.
son sometimes wants to use his or her real name and other times one or more persistently pseudonymous identities. The overarching problem is that most sites, including the three case studies examined earlier, do not support either standard, but it is my contention that these three problems would still be present if they did.

The first problem with current solutions is that they link identity in a "read-only" way. That is, none of the identities can actually affect one another; rather they are only aware of one another. This has negative implications for building trust due to the lack of credible punishments and thus does little to overcome the defection of the prisoners' dilemma and link present to future payoffs in order to create a coordination game. For example, even if eBay supported openID so that one could show others one's eBay score, this score would only be a signal of past cooperation. Because any transaction that happens outside eBay would not affect one's eBay score, the rational choice as a seller would still be to defect from shipping the item once it was paid for. Similarly, and from the Dick and Jane example, if Dick decides that Jane (on Twitter) is not a trusted friend anymore and decides to cancel the connection, the two may still be linked on LiveJournal as JaneEyre and RichardII. A system possessing read-write power would differ from current models of consolidation, because it would have the ability to change identity and reputation information across contexts rather than only extract this information for use elsewhere. In certain situations, however, read-only is the correct access permission: if one plays both WoW and Everquest (another MMORP), it would not make sense to increase one's WoW character level due to an increase in Everquest level. It would be useful, though, to pull in character and guild information from each game in order to prove skill as a new play-
er and encourage cooperation with the punishment that if one gets kicked out of a
guild in one game, this information would show up in the other. But in many con-
texts, read-only access leads to less effective trust, because if one identity is "spoiled"
by uncooperative behavior, one can always unlink it without consequence to the
others.

This problem is a subset of another large issue: openID and Social Graph fo-
cus almost exclusively on identity, rather than reputation, information. I say "almost"
because reputation can often be adduced from the identity information accessed
through these protocols, but doing so adds a cost to building trust. With openID, one
would need to use the site and username ownership proofs to actually go and visit
each site to make reputation extrapolations from each one. The same holds true for
Social Graph, at least in its capacity for linking a person's profiles. Reputation would
be easier to adduce through its ability to help visualize a person's connections across
the Internet—if connections are seen as implicit recommendations as they are on Face-
book—but this also requires costly interpretation. Again, this interpretation can have
value (as was shown through the WoW and Facebook case studies), but is not always
necessary or preferable for short-term and low-risk interaction. As it stands, neither
openID nor Social Graph does an effective job of broadcasting or linking reputation,
instead leaving this fundamental task up to the requesting user to interpret using the
identity information both standards make easier to find.

Adding this capability to the identity layer will make it simpler for requesters
to judge reputation. How it does so must be customizable based on the requester's
preferences. One way to customize is by using the identity layer to display a consoli-
dated reputation score about a person based on metrics that the requester provides.
For example, a user interested in trading fairness might define this score to take into
account signals such as an eBay feedback score of over 100, an email address from a
trusted domain (such as a known company), the knowledge that the seller lives in the
United States, and the information that the two have four friends in common (or even
one who has bought from this seller before). Someone asking a question about health
on a site such as Yahoo! Answers or ask.metafilter.com might set his computer to re-
turn only answers from those respondents who have answered a certain number of
questions in the past and had their answers rated well, and belong to the American
Medical Association or a hospital network on Facebook, a reliable signal of being a
doctor or nurse. Such information could be displayed with links to find out more, or
as a consolidated score that combines metrics and can be set up to process them by
the user's standards. The specific metrics are less important than the requirement that
users have choice.

Third, both openID and Social Graph follow a sort of "all-or-none" philoso-
phy and do not offer fine-grained identity and reputation information or access con-
trol. By "grain," I mean the "size" of a piece of information. Fine-grained identity in-
formation would be someone's address, while coarse-grained would be what country
he or she lives in. This information is meta-identity information, which also includes
statements such as "this person has been trustworthy, as measured by how often they
fulfill their agreement to ship an item once paid;" "this person has answered others'
questions four more times than others have answered his or hers;" and "this person
has five friends in common with you." Fine-grained access control, on the other
hand, is the ability to define that certain people or groups have access to differing levels of specificity, while coarse-grained would be having only the ability to either allow everyone or no one to see the information. Fine-grained access control is critical, because in certain contexts people want to reveal varying levels of information to different groups, something that neither openID nor Social Graph supports. An effective consolidated system would have to provide more fine-grained control so that users could decide how much identity and reputation information to broadcast, depending on the context and level of trust they perceive it requires, thus fulfilling Cameron's second law of identity: "minimal disclosure for constrained use...the solution that discloses the least amount of identifying information and best limits its use is the most stable long-term solution."\(^{202}\) Having fine-grained control over the level of meta-identity broadcasted is the best response to this imperative, but where this information comes from is a problem dealt with in the next section.

For a consolidated reputation system to work effectively while still protecting users' privacy, it must also give them fine-grained identity broadcasting control. This control should allow multiple identities, pieces of which are able to be separated and combined as needed. Microsoft's CardSpace system, to some extent, supports this choice, but as a primarily enterprise authentication solution that does not broadcast identity to other users, it is beyond the purview of this thesis.\(^{203}\) In the example of the doctor giving advice, this doctor may want to verify the specific hospital that she works at in some situations, but in others just prove that she works at a hospital. In other situations, the doctor may want to use an anonymous screen name to discuss radical politics on a message board where she does not want her postings associated
with her offline identity. A seller might maintain two pseudonyms on eBay and other auction sites, one of which he uses to sell antiques and therefore links to his Facebook account and offline business records to prove he is established. He may not want the other, which he uses to sell low-cost iPod accessories shipped directly from Hong Kong, associated with his antique business. In WoW, a player may want to give a prestigious guild access to her Facebook information as a sign of good faith that she will devote a lot of time to playing by giving up her right to separate game play from offline life (per the handicap principle), but the player may not want to broadcast this information to all characters in the game. Fine-grained identity broadcasting also includes the ability to decide what meta-level of information others have access to, as well as letting users define whether others can judge their reputation using an explicit score, implicit recommendations, or a performance-based interpretation. These types of fine-grained control (permissions and meta-level) are critical to linking identity in a way that still preserves the pseudonymous interactions that the Internet should continue to enable.

Current models of consolidation suffer from these problems of having only read-write capability, not being able to handle reputation information, and offering only coarse-grained control. The shortcomings of openID and Social Graph, for example, prevent them from adequately linking identity and reputation on the Internet while preserving user choice. They do, however, represent a step in the right direction and suggest the future of the Internet as supporting an identity layer that makes

* In other words, a requirement for the system is that it support multiple pseudonyms, which can be linked and unlinked depending on the signaling reliability needed for a certain context.
creating trust across contexts more efficient. My distributed oracle, I believe, is the next step in fulfilling this vision of the future.

3. The Distributed Oracle

Bob Blakley submits the need for an "identity oracle" that requesters can query for identity and meta-identity information, but this idea, while useful in enterprise contexts, is too centralized to be effective across the varying contexts that feed information about users. People keep their identity and reputation on many different websites, so forcing them to house it in a central data silo is unrealistic and inefficient. Instead, I posit that enlarging Greif's coalition, so that users are able to reliably signal to one another and build trust, requires a way to access this information across the distributed contexts in which it is now housed, like an expanded version of openID and Social Graph. I call the system that would enable this the distributed oracle. The distributed oracle is an agnostic set of protocols that connect websites, and allows them to share identity and reputation information without their having to renegotiate the terms of this connection every time. Websites can query the oracle, for example, on whether a user is likely to be from the United States as judged by the information on their Facebook profile or whether the user has an eBay feedback score of ten positive ratings or above. The oracle then requests this information from the appropriate context and returns the answer; how it does so is of little interest to the requesting website, as long as it gets the information it requests. The distributed oracle does not exist in any one "place;" instead, it floats between the contexts, acting as a
mediator for identity requests. Thus, it *consolidates* this information without *centralizing* it.

In contrast, Blakley's identity oracle takes the form of a single firm that stores and distributes identity information to requesters. This centralized reputation structure is potentially useful for building trust between a company and an individual, because the company will get risk-assessment information to decide whether or not to enter into the transaction and has the legal system as a credible punishment, which encourages the individual to cooperate. In addition, companies may require information such as credit scores and criminal records, which are best stored in centralized repositories. As we have seen, the type of identity and reputation information required to build trust among individual actors and groups already exists in many forms, but is locked up in contextual silos.

Transferring this information to a central identity oracle adds another layer of complexity and is inefficient. Instead, a distributed identity oracle, in the form of open protocols in the style of openID and Social Graph, but that addresses the problems identified above, is a more efficient solution to the problem. Identity and reputation information should be stored in its original context, so that it can be easily referenced, retrieved, and updated. Such distributed storage of the information would enable the fine-grained control of whether to broadcast reputation information (such as connections on Facebook) or gives others' access to the source (one's actual Facebook profile). Distributed storage also solves other operational problems. To use familiar examples, one's eBay feedback score, level and guild membership in WoW, and web of connections in Facebook are all constantly changing. Keeping them
synced with one central identity oracle would be structurally tricky and costly. Furthermore, some of the information does not lend itself to being exported and stored because it is dynamic and so firmly context-based. For example, the performance aspects of WoW reputation must be experienced within the game because they are so interpretation-based. Activity or chat logs could be theoretically extracted from the game (adding to The Armory's listings of items, level, and guild membership), but it is far more efficient to view them within context.*

A distributed oracle would act as a reputation courier, fulfilling requesters' demands for information about others only with their consent. The beauty of this is threefold. First, information is only stored in places that users already trust. In other words, if someone is using Facebook, it is assumed that he trusts Facebook to keep his data safe and only give it out with approval. The distributed oracle follows these rules, taking the data only in the form it is authorized to and not requiring users to give up the full data set to another storage context. Second, the information will always be as up to date as users want it to be, because it is accessed from the place where users already put it, not some third party. Third, the system is flexible. It supports the fine-grained access and broadcast control that was identified above as sorely lacking in openID and Social Graph. Because information retrieval is done in the form of "calls" to various silos, it supports meta-identity information (age range), specific information (exact age), or full access (exact age and the profile information that

* This would mean needing to actually let others get into the game to watch a character's performance. Because this would be time-consuming and most likely disallowed by Blizzard, proving ownership of a character and being able to distribute level, guild membership, and items ownership would usually be acceptable.
"proves" it.). The ability to set permissions ensures that users still have control over their privacy and a choice over the strength of signal they want to send. Giving access to more information represents increased pre-play communication and is a stronger identity signal by the handicap principle that makes others' reputation judgments easier, while giving less will suffice as a weaker signal in some contexts and protect more privacy.

Furthermore, to work correctly, the distributed oracle must have the "read-write" capability among contexts that was touched on above. This power not only ensures that requesters have more reputation information about others to use, but that they know they have a credible punishment strategy beyond the current context if the relationship goes sour. Recalling the section on game theory, this turns a series of prisoners' dilemmas into one coordination super-game by linking present and future payoffs, because a negative reputation harms a player. This is what eBay accomplishes with its feedback score, but in its case, the coordination super-game is only applicable to transactions on the eBay site. A distributed oracle that links silos extends the benefits of this super-game (namely that players have a chance of reaching the Pareto superior strategy set of Cooperate-Cooperate) to all contexts to which it is attached, thus increasing the size of the coalition in which reputation matters. Again, users and site operators must have control over what silos are able to read and write to one another. It would be difficult to convince websites to let the distributed oracle write to the identity and reputation information they store, as companies are protective of letting others into their systems. As such, read-write capability might not be a
practical demand, but nonetheless one that a consolidated system should strive towards.

An example will permit deeper understanding. A seller with a high feedback score on eBay may want to start selling on uBid.com (another auction site specializing in large-scale inventory liquidation). Starting off, the seller may be subject to the punishment of lower prices until he or she builds up a good feedback score on uBid. Even if sellers could prove their feedback score on eBay, uBid buyers would have no recourse to punish this score if a seller failed to deliver on uBid. It is plausible that sellers who want to unload sub-par merchandise using their eBay feedback score, but not risk harming this reputation, would go to uBid. Read-write capability solves this problem. Using the distributed oracle, sellers on eBay and uBid would be able to set their feedback scores to display in both contexts as well as mutually update, thus reliably signaling by giving up one's immunity to punishment. * Read-write capability extends the coordination super-game to operate between now connected contexts by linking punishment scores. To accomplish this, new "tags" are needed to enable ratings that meaningfully describe the outcome (whether positive or negative) of relationships in contexts as diverse as gaming, blogs, SNSs, and trading sites. ** The dis-

* Of course, this does not guarantee that a seller does not plan to continue using either forum. It also assumes that eBay would let uBid, a competitor, update its feedback scores.

** These tags could include things like "good warrior," "honest trader," and "trustworthy source." How they would be implemented is up to each site that chooses to use the distributed oracle. The OASIS Open Reputation Management Systems Technical Committee is working to define what these tags might looks like and how to share them among websites. Because the group has not met yet, it is too soon to tell what their solution might look like (Cover, 2008).
tributed oracle reads these tags in order to port the information they describe to other contexts or into a consolidated score.

Even though the distributed oracle does not rely on a centralized identity and reputation provider, it could be configured to feed information to a centralized display. Websites serving this purpose would most likely emerge and offer the distributed oracle's information gathering power to users who desired a way to display their contextual reputations in one place. Repatoo.com offers to centralize one's auction site ratings as a signal to potential buyers, but without read-write capability it is ineffectual. FriendFeed.com aggregates one's online activity into one page for others to see, but to pull in data from many services, a user does not even need a password to prove that the data belongs to him or her, making it a weak signal. openID providers such as claimID.com and myOpenID.com have emerged to fill a similar role, but they are subject to the limited usefulness of openID for only proving URL ownership. The distributed oracle would add reputation to the current identity capability of these classes of sites and offer a way to prove one's various ratings and presences across the Internet, while also enabling a read-write option. Permissions of what centralized displays have access to would work in the same way as described in the fine-grained control discussion above. Such a system is not immune to fraud, but with its increased number of linkages and heightened identity retrieval powers, it adds more information and signals for users to consult in making decisions of whether to cooperate or defect, answer someone else's question or ignore it, and trust someone's answer or not.
It is not guaranteed that distributed consolidation will win out over centralized consolidation. The company or group that first develops a widely used identity layer for the Internet will end up with a lot of power. Controlling users' identities and reputations online is a goal that many companies (particularly Oracle, Computer Associates, BEA, IBM, and Microsoft) are actively pursuing, because of the substantial resource that doing so would give them. Being able to sell the right to query this information to other websites and companies would produce sizable profits. The potential for ill-use is also obvious, partly explaining why no company has been able to capture this market so far. One can imagine the potential privacy abuses if this information was sold to marketers, for example. For this reason, Blakley's centralized identity oracle that was encountered earlier is not the correct solution to the problem. Keeping the oracle distributed reduces the potential for one company to exploit the information to which it offers access. How a company or other group could bring a centralized or distributed oracle to market is a question for another thesis, but suffice it to say that it is one embroiled with politics and power struggles as many interested parties (corporate, governmental, and non-profit) jockey to control who you are online.

The distributed oracle enables an identity and reputation layer on the Internet to connect any networked component (whether website, program, or mobile device) by a set of protocols that permits consolidated sharing and retrieval of identity information in order to make reputation judgments. As such, it is similar to the World Wide Web, a set of protocols that links information dispersed among servers across the world into an easy to use interface. Since the advent of the Web, many programs
have been built to access and harness the information it makes available (web browsers and search engines for example). Similarly, for it to be successful, tools would have to be built to make the distributed oracle (which is no more than a set of protocols extending the type of information sharing and ownership proving that openID and Social Graph make possible) useful. These might take the form of websites that gather one's dispersed identities to make them easily viewable in one place, algorithms that create a rating number to describe a person based on individually defined metrics, or programs that keep track of one's Internet wide reputation.

I am not suggesting that the programming required to create the distributed oracle is simple. Open protocols such as this, however, are not without precedent. openID and Social Graph, as well as the large number of discussions and conferences devoted to online identity, reveal that the industry realizes the importance of improving reputation reporting so that individuals can engage in useful interactions online.* The distributed oracle will free identity and reputation data from the silos in which they now reside and create the potential for cheaper, temporary ties in niche contexts that currently require prohibitively high transaction costs to be useful. Websites will then be able to use the trust that is created to help support whatever purpose they choose, whether this be enabling more trades, building networks, or getting users to play a game. These large potential benefits will, I believe, offset the costs of designing the system, just as they have for openID and Social Graph.

* See, for example The Data Portability Project (DataPortability, 2008).
4. Problems

The distributed oracle and the consolidated reputation and identity system it creates are not without their own problems. First, it is a potentially difficult set of protocols to design, due in part to the fine-grained control it must give to identity requesters and broadcasters. Issues would also likely arise in regards to who controls the standard, as companies clash over specifics that impact their businesses. As this thesis focuses primarily on the benefits that such a system has for interaction on the Internet, technical design challenges are beyond its scope and are best taken up by industry or other academics. The two main theoretical problems, though, have to do with privacy and the potential to defraud the distributed oracle.

Securing the right to privacy while enabling reputation consolidation is an important question, though one that has been addressed above. This problem is largely solved by giving users the fine-grained choice that I discussed in the previous section. Nonetheless, if abused, a consolidated system like the distributed oracle could let some people acquire more information about others than those others think they are revealing. Preventing this outcome would largely lie in the hands of the application developers who used the oracle's protocols, but following Cameron's first law that "technical identity systems must only reveal information identifying a user with the user's consent" is a good starting point. For example, sites tagging links with XFN or FOAF meta-data (which Social Graph looks for) should give users a choice of whether to turn this tagging on or off. If anything, the distributed oracle might increase privacy controls by giving users more fine-grained choice, rather than the on/
off option they have now, over what they reveal in different contexts. Nonetheless, we would be wise to heed Danah Boyd's warning against exposing users' information to the point that they "become hermits to cope with exposure," \(^{210}\) thus acting against the goal for which the distributed oracle is proposed—

*increase* social interaction.

Privacy is a critical concern, but one is solved by giving the choice of how much identity and reputation to broadcast to the user, the entity for whom this system is ultimately created.

The distributed identity, while it does much to curb it, will not eliminate fraud. There is still the ever-present issue that past action does not guarantee future action. Linking reputation and punishments between silos helps curb this problem by increasing the number of available signals and linking payoffs between single prisoners' dilemmas to create the coordination super-game. Still, there is always the possibility that someone trying to game the system will build up a good reputation among all the linked sites only to rip-off the next person at large enough stakes that the fraudster does not mind spoiling his or her entire entire reputation associated with that pseudonym, just like the penny e-book fraud discussed in the eBay section. As discussed in the game theory section, it is nearly impossible to prove that one is going to "stick around" and keep playing the game in all of the contexts from which one's reputation comes, especially if it is not attached to one's real name. Hooking Facebook up to the distributed oracle signals future participation, because any reputational hit is potentially broadcast to one's real offline friends. Nonetheless, as this reputation becomes even more important and useful in other contexts, creative ways will likely be found to defraud it. One blog predicts that paying for Facebook friends will become a com-
mon occurrence in the near term, and one can imagine complex schemes to connect with certain key trusted people on the site in order to falsely signal trustworthiness. Moreover, fraudsters will be able to take advantage of the fact that evaluators will have to interpret a wide variety of identity and reputation sources, leading to the potential to pull in information from illegitimate contexts to legitimate ones and trick users who do not know better.

The final problem that has the potential to derail the distributed oracle is at the service provider level. I predict a free rider dilemma among websites who, realizing they will be able to offer their users reputation information from other contexts using the distributed oracle, do not contribute their own identity and reputation data for others to use. This will lead to a mass defection in the form of sites keeping their data closed off, because defection (not sharing identity information with the distributed oracle) is the rational choice whether or not others cooperate (share identity information) or defect. For example, if a new auction site is created, why would they bother to create their own difficult to design reputation system when they can just pull in eBay's data? eBay, in response, may disconnect themselves from the distributed oracle by making profile pages private and not providing a pipeline in for the oracle to use, or by only making their reputation data available for a cost. As a result, in classic prisoners' dilemma form, no identity or reputation information will be shared in the equivalent of a no-trade equilibrium. To prevent this free-riding, the distributed oracle

* This effect is heightened by the asymmetry of the reputation information different sites provide. That is, eBay's feedback ratings are significantly more valuable than the ratings of a small would-be competitor, making the benefits that this young site would get from using the distributed oracle larger than those eBay would receive.
oracle might require that only websites who share identity information are able to access identity data from other contexts. Websites, seeing the boon to their purposes from the trust created by easier reputational evaluation, will most likely choose to open up their data in exchange for getting access to others' data. This data will be such an asset to trust creation that any website that requires interaction among its users will not be able to risk losing users to sites better equipped to engender trust with the information the distributed oracle provides.

Despite its problems, the distributed oracle is a better option for trust-supporting than the siloed systems that are currently the norm. More information to use for reputation evaluation equates to more pre-play communication, which can lower players' risk aversion, even if the information is not fool-proof. As they must now, users will have to separate conventional and reliable signals and weigh risks when evaluating others' identities and reputations. By having access to more data, and in potentially varying forms (explicit, implicit, and performance-based) users will be able to make more informed decisions that lead to trust in widely varying contexts. The distributed oracle that I have described accomplishes this more effectively than current models of consolidation such as openID and Social Graph, because it offers the read-write power that makes reputation a credible reward and punishment among contexts. By also supporting fine-grained access and broadcast control, it addresses many of the privacy issues, as long as websites that use its protocols pass this choice on to their users.
5. Conclusions

Under the current model, identity and reputation are siloed within walled gardens from which they cannot escape. If one builds up a reputation in one, it is difficult if not impossible to carry it into other contexts. This creates inefficiencies, because users must build up reputations on every service they use, an activity that takes time and effort. Furthermore, users have only this siloed identity information to use in judging reputation, a restriction that makes it easier for others to broadcast fraudulent identity information. These problems make it more costly to build trust between users. As we have seen, some level of trust is needed to support the purpose of any site that requires users to interact with one another. Without this trust, relationships break down as users defect by free riding on others' contributions and losing confidence in others' abilities, which results in fewer social interactions. While connections may be able to persist within sites that maintain effective reputation systems, they cannot escape from these boundaries, leading to a model in which users are wary of trying new services.

The distributed oracle offers a way to connect previously siloed identity and reputation information into a consolidated system that makes it easy for users to query information about others, even if they are not established within a certain context, thus increasing the size of Greif's coalition in order to strengthen the signals that overcome information asymmetry. In this way, it turns identity and reputation on the Internet into an ecosystem, where everything is able to affect everything else. Just as the warming arctic waters affect weather in the tropics, with these protocols one's
Facebook reputation affects one's eBay seller rating. Unlike the natural world, however, users must have choice over how wide-ranging they want to make their identity ecosystem, as well as which components will make up its parts. Thus, rather than allowing a global Internet reputation, a user may choose privacy settings that result in an online reputation more analogous to the internal ecosystem of a rainforest, where certain identity information travels outside the boundaries while other information is contained. This choice affects what types of signals users send to others, and with enough strong signals (the number of which is increased by the distributed oracle being able to access information across contexts), reputation becomes harder to defraud and more likely to convert a series of prisoners' dilemmas into a coordination super-game. Websites or other networked applications, as the theory predicts, can then draw on the trust these signals create to support whatever purpose they choose. Figure 10 compares the old model of identity and reputation and its effects—in place on sites like eBay, WoW, and Facebook—with the new model that the distributed oracle would enable.
This type of system would elegantly follow the arc of the Internet's development. From unconnected networks in the seventies, to the Internet that linked these networks in the eighties, through the World Wide Web that made information retrieval far easier in the nineties, and finally to the social-focused Web 2.0 trend of the two-thousands, long-distance networked connection has become increasingly consolidated in such a way that makes it easier for people to connect and find information.
An identity layer will continue this trend by supporting social interaction between website contexts that are still largely separate, similar to the networks of the seventies until the Internet joined them together. The protocols that the distributed oracle would define are one way to further connect those who increasingly use the Internet in all aspects of their lives. They would reduce the time and effort involved in finding an informed expert to answer a question, a player to quest with in WoW, a connection to introduce a potential job lead, or a seller from whom to buy vintage Pez dispensers. Without identity and reputation consolidation, I fear, the Internet will continue to be fragmented in such a way that makes useful social interaction too costly to undertake. The resulting inefficiencies will result in a connective medium that further spirals towards fraud and deception and results in defection among users who are unable to identify or judge one another. A system like the distributed oracle, I believe, is necessary to stem this decline.


**CONCLUSION**

Despite connecting people who may never see one another face-to-face, sites on the Internet have developed ways for their users to judge one another's identities and reputations. Given the theory and evidence presented above, the basic fact that perfect strangers are able to build trust online that supports their interactions in contexts as diverse as auctions, games, and networking should not be surprising. But this was not always possible online, and early applications of the Internet lacked organized reputation systems that made it difficult for them to grow beyond their small, devoted user bases. With the advent of the World Wide Web, and the increased number of users who flocked to its simple system of URLs and hyperlinks, websites found it necessary to develop mechanisms to support these users' social interactions. As time went on, designers developed some of the systems examined in the case studies, making it possible for people to engage in short-term connections online. Now, in contrast to the 1993 *New Yorker* cartoon, it might be said that "On the Internet, everybody knows you're a dog."

The theory presented a model that shows how reputation actually supports the trust needed for short-term social interactions. Trust, as will be recalled, is the prediction of one actor that another will do something previously agreed upon, whether implicitly or explicitly, even when this may not be in the trusted actor's short-term selfish interest. Trustworthiness is one kind of reputation, which makes up the adjectives (such as dependability or skill at a certain act) that describe a person. Identity, on the other hand, includes the nouns that describe a person—things like name, loca-
tion, and history of past action. When people interact with one another, as I have shown, it may be difficult to broadcast or ascertain all the necessary identity and reputation information needed to support the interaction. This problem creates the potential for an information asymmetry of the kind that Akerlof described. Signals are one way to overcome the asymmetry, but they run into the complication of how to determine whether or not they are reliable. Reputation and identity information are signals that actors can send and use, and with a reliable system in place to support these signals online, actors can discern the information they need to build trust. Furthermore, when actors can determine one another's reputations, this becomes a credible reward and punishment system. Reputation links present and future payoffs in order to connect a series of single-shot prisoners' dilemma games into the coordination super-game, which results in cooperation as long as trust checks players' risk aversion. Thus, an identity and reputation system helps players overcome their distrust of one another so that they can enter into useful social interactions.

The three case studies applied this theory to real-world examples. eBay supports reputation through a relatively simple and explicit system: trading agents rate one another after undertaking a transaction, and the cumulative score resulting from this feedback is displayed next to each user's pseudonym. Clicking on a username brings up a more detailed profile page that lists positive, neutral, and negative feedback over various periods of time. Sellers' ratings have more effect on their success on the site than buyers', and research has shown that a better feedback rating results in a higher probability of sale as well as a higher sale price. The simplicity of the system is both a blessing and a curse. On the one hand, eBay has enjoyed a high usage
rate, because it drastically lowers the cost of evaluating reputation. On the other, its reputation system has proven vulnerable to fraud and the age-old dictum that "past performance doesn't guarantee future results."

The World of Warcraft, as a result of the massive life-mimicking virtual world in which players interact, uses a reputation system that is largely performance-based. In this sense, it mirrors offline life, in which players must judge one another using signals garnered from interaction, though players do not have the benefit of face-to-face contact. WoW does have explicit reputation measures: level and guild membership are often a good indicator of player skill, and players use this information to evaluate grouping and other decisions that require trust. But these explicit measures are subject to fraud, most notably by players purchasing pre-built characters. Thus, most reputation broadcasting takes place through performance. Players display rare items and skills they have acquired, brag about familiarity with arcane facets of the game and little known combat tactics that could only be figured out after months of play, and prove their worth through solid performance in battle. Performance-based measures are effective signals of reputation not only because they are difficult to fake, but also because they enable ample interpretation on the part of the evaluator. This latter property allows players to judge others on the metrics that they find important and to use the instinct that is largely absent when relying on pre-packaged explicit measures. However, these performance-based measures are costly to evaluate. Thus, there is a tradeoff in terms of complexity and cost, amount of information and required interpretation.
Finally, Facebook uses an implicit model of reputation that is primarily based on the recommendation suggested by a connection with another user. While users have other signals at their disposal, such as information included on the profile page and membership in various restricted networks, the most reliable signal stems from users' friends. If an unknown user has many friends, especially friends with whom one is also connected, this is a signal of truthful identity. On Facebook, reputation means a different thing than it does in the other case studies, because it is so intimately connected to offline life. Users may consult Facebook to obtain reputation information about unknowns to use in other contexts ("Does this person's profile make them seem responsible?"), but judgments are primarily made about users' identities. Trust is predominately a measure of how truthful others are about their identities, which are expected to reflect their "real," offline selves. This is in contrast to eBay and WoW, in which pseudonyms, though persistent, are not expected to reflect a user's real name. As might be inferred, identity evaluation on Facebook requires a great deal of interpretation, but observational evidence suggests this does not stop users from friending one another.

These case studies largely supported the theoretical claims I made, but each revealed that there is room for improvement in how identity and reputation are signaled online. My distributed oracle strengthens the dependability of identity and reputation information by linking the siloed contexts in which it currently resides. Doing so results in the benefits, which Greif described, of enlarging the coalition within which reputation matters. Emerging standards such as openID and Social Graph are the first step in doing so, but they have significant limitations and do not strongly
enough link identity and reputation information. The distributed oracle, on the other hand, allows users to link reputations according to their preferences; as such, it preserves user privacy and choice while also supporting more reliable signals. In this capacity, I submit that it represents a natural next step in fulfilling the theoretical claims and suggestions I have made, and is also a practical extension of current attempts to design an identity layer for the Internet. Whether such a system is feasible or how it will actually affect social interaction online remains to be seen, but I believe that current trends suggest that we will have an answer soon enough.

There is ample room for further research. Far more empirical research is needed to measure the effectiveness of the case studies' reputation systems in forcing users to consider the future. Ethnographic-style research might be used to describe how these users interpret signals on these services in order to better understand how users approach social interaction and come to trust one another. More technical research is also needed on how the protocols of the distributed oracle would actually look. What tags would be needed to allow ratings that could easily travel between contexts? How easy would it be to aggregate the opinions of trusted individuals into the "wisdom of the crowds" model that is popular now? Furthermore, deeper thinking would be required to understand whether the average user would actually link their dispersed identities, and whether weighing the potential privacy issues would be easy enough to understand. Empirical work is also needed to assess whether users would actually rate one another across contexts, as well as whether the consolidated reputation system would succumb to free rider problems. Researchers might also examine how well reputation reinforced trust supports social capital and online com-
munities. I believe that well-designed studies would easily engage these questions and further inform the theories that I have discussed here.

The Internet is a constantly evolving place. As so much economic and social activity has shifted to operate within its sphere, questions of how to continue to support these activities take on increasing importance. Many take for granted how easy it is to Google the answer to a question, find a friend's phone number, and order a hard-to-find CD. These pursuits will not necessarily continue to be possible. Fear rises in tandem with news reports of increased fraud on the Internet, and without a reliable identity and reputation system, I believe, users will slowly lose the ability to evaluate those with whom they deal and thus become too frightened to act online. Because systems that are in place to stem this decline may not have been fully tested yet, deeper thinking is needed about the strength of the trust and types of interactions they enable. New models are needed to replace them. Like clippers tossed about on the rough seas of change, siloed identity and reputation will falter against the harsh winds of mistrust if they do not evolve by forming an interwoven system held together by the distributed oracle. Without this change, the Internet may be left a useless stage for constructive social activity.
ENDNOTES

1. Paisley, 2007
2. Poole et al., 1999, p. 52; Gillies and Cailliau, 2000, p. 12
3. O'Reilly, 2006
4. Cameron, 2006
5. Poole et al., 1999, p. 103
7. Donath, 1999, p. 35
8. Donath, 1999, p. 40
9. Poole et al., 1999, p. 127; Rheingold, 2000
10. Berners-lee, 2000
12. Steiner, 1993
13. Confucious, 2004
15. Bowles and Gintis, 2000
17. Putnam et al., 1994, p. 171
18. Lin, 2001, p. 147
20. Gambetta, 1990 O'Neill concurs with the subjective nature of trust, positing that it is difficult to measure, and thus that discussions about it are best kept theoretical, and, I would add, qualitative (O'Neill, 2002).
21. Abdul-Rahman and Hailes, 2000, p. 3; Bacharach and Gambetta, 2001
22. Abdul-Rahman and Hailes, 2000, p. 2; Masum and Zhang, 2004
25. Putnam, 2000, p. 21
27. Piper et al., 2005
29. For more information on authentication, see Windley, 2005; Mansell and Collins, 2007.
30. Blakley, 2006
31. Abdul-Rahman and Hailes, 2000, p. 3
32. Dellarocas, 2003
33. Granovetter, 1973; Granovetter, 1983
34. Putnam, 2000, p. 23
35. Greif, 1989; Greif et al., 1994
36. Akerlof, 1970
37. Greif, 1989, p. 859
38. Greif, 1989, p. 868
39. Greif, 1989, p. 881
40. Dellarocas, 2003, p. 1410
41. Piper et al., 2005, p. 94
42. Dunbar, 1992
43. Masum and Zhang, 2004
44. Ghosh, 2007; Chamlee-Wright and Myers, 2008
45. Misztal, 1996, p. 31
46. Donath, 1999, p. 228
47. Kullok, 1999
48. Zahavi, 1975; Zahavi, 1977
49. Donath, 2007a, p. 3
50. Donath, 2007a, p. 3; Donath, 2008, p. 8
51. Donath, 2007a, p. 11
52. Akerlof, 1970
54. Spence, 1973
55. Spence, 1973, p. 358
56. Clippinger, 2007, p. 158
57. Donath, 2007b
58. Donath, 2008, p. 9
59. Donath and Boyd, 2004, p. 73
60. Donath and Boyd, 2004, p. 76
61. Donath, 1999, p. 25
63. Basu, 2003, p. 16
64. Basu, 2003, p. 18
66. Hardin, 1968
68. Dellarocas, 2001, p. 1
69. Friedman and Resnick, 2001, pp. 173-174
70. Miller et al., 2002, p. 1
71. Axelrod, 1984, p. 126
72. Cooper, 1999 has excellent research on the topic of pre-play communication in non-cooperative games.
73. Bolton et al., 2003, p. 23
74. Dellarocas, 2003, p. 1414
75. Dellarocas, 2003, p. 1415  See also Resnick et al., 2006, p. 82.
76. Basu, 2003, p. 34
77. Mailath and Samuelson, 2006, p. 2
78. Cooper, 1999, p. 7
79. Cooper, 1999, p. 11
80. Cooper, 1999, p. 11
81. This paragraph is taken from Rosen, 2007.
82. Axelrod, 1984, p. 139
83. Cabral, 2005, pp. 16-17
84. I am indebted to Professor Gil Skillman for especially helpful conversation on these points.
85. Bolton et al., 2003
86. Bolton et al., 2003, p. 20
87. invisiblog, 2008
88. Baron, 2002, p. 2
89. eBay, 2008b
90. Waring, 2008; Bolton et al., 2003, p. 3
91. Resnick and Zeckhauser, 2002, p. 9
92. Bolton et al., 2003
93. Baron, 2002, p. 8
94. Melnik and Alm, 2005, p. 310
95. Resnick and Zeckhauser, 2002, p. 11
96. eBay, 2008d
97. eBay, 2008d
98. Fehd, 2008
99. Waldmeir, 2008
100. eBay, 2008a
See, for example, Williams et al., 2006, p. 352.

Ducheneaut et al., 2006, p. 409

"VanGoghPower-GamerAtLarge", 2007

Yee, 2007, p. 11

Nardi and Harris, 2006, p. 149

Ducheneaut et al., 2006, p. 413

Ducheneaut et al., 2006, p. 413

Heaniis, 2008

BlizzardEntertainment, 2008

Christopher, 2006

Hsueh-hua Chen and Duh, 2007, p. 23

Ducheneaut et al., 2006, p. 414

See, for example, Putnam, 2000 and Lin, 2001 for more discussion of social capital.

WoWJutsu, 2008

Ducheneaut et al., 2007, p. 839

Ducheneaut et al., 2007, p. 839

Manninen and Kujanpaa, 2007, p. 26

Nardi and Harris, 2006, p. 155

Manninen and Kujanpaa, 2007, p. 6

Nardi and Harris, 2006, p. 151

Dibbel, 2006

For an excellent short documentary on the practice, see (jingejinge, 2006).

2008b

2008a

wikiHow

Ducheneaut et al., 2006, p. 415
153. Ducheneaut et al., 2007, p. 846

154. This language comes from the email that Facebook sends out when requesting friend confirmation.


156. Boyd and Ellison, 2007, pp. 2. See also Alessandro and Ralph, 2006, p. 2; Golder et al., 2007, p. 2.


158. Stutzman, 2006, p. 3; Golder et al., 2007, p. 2

159. DiMicco and Millen, 2007, p. 383

160. Lampe et al., 2007, p. 437


162. Donath and Boyd, 2004, p. 73

163. Lampe et al., 2007, p. 437

164. Donath and Boyd, 2004, p. 73


166. Stutzman, 2006, p. 5

167. Donath, 2007b

168. Stutzman, 2006, p. 6

169. Acquisti and Gross, 2006, p. 3

170. Donath and Boyd, 2004, p. 81

171. Donath, 2007a, p. 4

172. Lampe et al., 2007, p. 440


174. Acquisti and Gross, 2006, pp. 8-9

175. Facebook, 2007

176. Finklestein, 2008
203. Microsoft, 2007
204. Blakley, 2006
205. Blakley, 2006; Blakley, 2006
206. Repatoo, 2008
207. FriendFeed, 2008. For more discussion of FriendFeed.com, as well as some of the problems it raises in regards to centralization and portability, see Arrington, 2008b; Le Meur, 2008.
208. claimID, 2008; myOpenID, 2008
209. Cameron, 2006
211. Xobni, 2007
212. Surowiecki, 2005
WORKS CITED


Characters in MMOGs. Int. Journal of Business Science and Applied
Management, 2(1).

Cheltenham, UK: Edward Elgar Pub.

Monday, 9(7).


content/WindowsCardspaceHome.aspx

University, Faculty Research Working Papers Series.


/www.myopenid.com/

Nardi, B. & Harris, J. (2006). Strangers and Friends: Collaborative Play in World of
Warcraft. Paper presented at the 20th anniversary conference on Computer
supported cooperative work, Banff, Canada.

Retrieved 1/17/08, http://thelede.blogs.nytimes.com/2008/01/02/prankster-
playing-bhuttos-son-on-facebook-fools-news-outlets/
index.html?ex=1357275600&en=7743652a25cccba2&ei=5088&partner=rssnyt
&emc=rss


radar.oreilly.com/archives/2008/02/social-graph-visibility-akin-t.html


