



Overview

Just today, a stranger came to my door claiming he was here to unclog a bathroom drain. I let him into my house without verifying his identity, and not only did he repair the drain, he also took off his shoes so he wouldn't track mud on my floors. When he was done, I gave him a piece of paper that asked my bank to give him some money. He accepted it without a second glance. At no point did he attempt to take my possessions, and at no point did I attempt the same of him. In fact, neither of us worried that the other would. My wife was also home, but it never occurred to me that he was a sexual rival and I should therefore kill him.

Also today, I passed several strangers on the street without any of them attacking me. I bought food from a grocery store, not at all concerned that it might be unfit for human consumption. I locked my front door, but didn't spare a moment's worry at how easy it would be for someone to smash my window in. Even people driving cars, large murderous instruments that could crush me like a bug, didn't scare me.

Most amazingly, this worked without much overt security. I don't carry a gun for self-defense, nor do I wear body armor. I don't use a home burglar alarm. I don't test my food for poison. I don't even engage in conspicuous displays of physical prowess to intimidate other people I encounter.

It's what we call "trust." Actually, it's what we call "civilization."

All complex ecosystems, whether they are biological ecosystems like the human body, natural ecosystems like a rain forest, social ecosystems like an open-air market, or socio-technical ecosystems like the global financial system or the Internet, are deeply interlinked. Individual units within those ecosystems are interdependent, each doing its part and relying on the other units to do their parts as well. This is neither rare nor difficult, and complex ecosystems abound.

At the same time, all complex ecosystems contain parasites. Within every interdependent system, there are individuals who try to subvert the system to their own ends. These could be tapeworms in our digestive tracts, thieves in a bazaar, robbers disguised as plumbers, spammers on the Internet, or companies that move their profits offshore to evade taxes.

Within complex systems, there is a fundamental tension between what I'm going to call cooperating, or acting in the group interest; and what I'll call defecting, or acting against the group interest and instead in one's own self-interest. Political philosophers have recognized this antinomy since Plato. We might individually want each other's stuff, but we're collectively better off if everyone respects property rights and no one steals. We might individually want to reap the benefits of government without having to pay for them, but we're collectively better off if everyone pays taxes. Every country might want to be able to do whatever it wants, but the world is better off with international agreements, treaties, and organizations. In general, we're collectively better off if society limits individual behavior, and we'd each be better off if those limits didn't apply to us individually. That doesn't work, of course, and most of us recognize this. Most of the time, we realize that it is in our self-interest to act in the group interest. But because parasites will always exist—because some of us steal, don't pay our taxes, ignore international agreements, or ignore limits on our behavior—we also need security.

Society runs on trust. We all need to trust that the random people we interact with will cooperate. Not trust completely, not trust blindly, but be reasonably sure (whatever that means) that our trust is well-founded and they will be trustworthy in return (whatever *that* means). This is vital. If the number of parasites gets too large, if too many people steal or too many people don't pay their taxes, society no longer works. It doesn't work both because there is so much theft that people can't be secure in their property, and because even the honest become suspicious of everyone else. More importantly, it doesn't work because the social contract breaks down: society is no longer seen as providing the required benefits. Trust is largely habit, and when there's not enough trust to be had, people stop trusting each other.

The devil is in the details. In all societies, for example, there are instances where property is legitimately taken from one person and given to another: taxes, fines, fees, confiscation of contraband, theft by a legitimate but despised ruler, etc. And a societal norm like "everyone pays his or her taxes" is distinct from any discussion about what sort of tax code is fair. But while we might disagree

about the extent of the norms we subject ourselves to—that's what politics is all about—we're collectively better off if we all follow them.

Of course, it's actually more complicated than that. A person might decide to break the norms, not for selfish parasitical reasons, but because his moral compass tells him to. He might help escaped slaves flee into Canada because slavery is wrong. He might refuse to pay taxes because he disagrees with what his government is spending his money on. He might help laboratory animals escape because he believes animal testing is wrong. He might shoot a doctor who performs abortions because he believes abortion is wrong. And so on.

Sometimes we decide a norm breaker did the right thing. Sometimes we decide that he did the wrong thing. Sometimes there's consensus, and sometimes we disagree. And sometimes those who dare to defy the group norm become catalysts for social change. Norm breakers rioted against the police raids of the Stonewall Inn in New York in 1969, at the beginning of the gay rights movement. Norm breakers hid and saved the lives of Jews in World War II Europe, organized the Civil Rights bus protests in the American South, and assembled in unlawful protest at Tiananmen Square. When the group norm is later deemed immoral, history may call those who refused to follow it heroes.

In 2008, the U.S. real estate industry collapsed, almost taking the global economy with it. The causes of the disaster are complex, but were in a large part caused by financial institutions and their employees subverting financial systems to their own ends. They wrote mortgages to homeowners who couldn't afford them, and then repackaged and resold those mortgages in ways that intentionally hid real risk. Financial analysts, who made money rating these bonds, gave them high ratings to ensure repeat rating business.

This is an example of a failure of trust: a limited number of people were able to use the global financial system for their own personal gain. That sort of thing isn't supposed to happen. But it did happen. And it will happen again if society doesn't get better at both trust and security.

Failures in trust have become global problems:

- The Internet brings amazing benefits to those who have access to it, but it also brings with it new forms of fraud. Impersonation fraud—now called identity theft—is both easier and more profitable than it was pre-Internet. Spam continues to undermine the usability of e-mail. Social networking sites deliberately make it hard for people to effectively manage their own privacy. And antagonistic behavior threatens almost every Internet community.

- Globalization has improved the lives of people in many countries, but with it came an increased threat of global terrorism. The terrorist attacks of 9/11 were a failure of trust, and so were the government overreactions in the decade following.
- The financial network allows anyone to do business with anyone else around the world; but easily hacked financial accounts mean there is enormous profit in fraudulent transactions, and easily hacked computer databases mean there is also a global market in (terrifyingly cheap) stolen credit card numbers and personal dossiers to enable those fraudulent transactions.
- Goods and services are now supplied worldwide at much lower cost, but with this change comes tainted foods, unsafe children's toys, and the outsourcing of data processing to countries with different laws.
- Global production also means more production, but with it comes environmental pollution. If a company discharges lead into the atmosphere—or chlorofluorocarbons, or nitrogen oxides, or carbon dioxide—that company gets all the benefit of cheaper production costs, but the environmental cost falls on everybody else on the planet.

And it's not just global problems, of course. Narrower failures in trust are so numerous as to defy listing. Here are just a few examples:

- In 2009–2010, officials of Bell, California, effectively looted the city's treasury, awarding themselves unusually high salaries, often for part-time work.
- Some early online games, such as Star Wars Galaxy Quest, collapsed due to internal cheating.
- The senior executives at companies such as WorldCom, Enron, and Adelphia inflated their companies' stock prices through fraudulent accounting practices, awarding themselves huge bonuses but destroying the companies in the process.

What ties all these examples together is that the interest of society was in conflict with the interests of certain individuals within society. Society had some normative behaviors, but failed to ensure that enough people cooperated and

followed those behaviors. Instead, the defectors within the group became too large or too powerful or too successful, and ruined it for everyone.



This book is about trust. Specifically, it's about trust within a group. It's important that defectors not take advantage of the group, but it's also important for everyone in the group to trust that defectors won't take advantage.

“Trust” is a complex concept, and has a lot of flavors of meaning. Sociologist Piotr Sztompka wrote that “trust is a bet about the future contingent actions of others.” Political science professor Russell Hardin wrote: “Trust involves giving discretion to another to affect one’s interests.” These definitions focus on trust between individuals and, by extension, their trustworthiness.¹

When we trust people, we can either trust their intentions or their actions. The first is more intimate. When we say we trust a friend, that trust isn't tied to any particular thing he's doing. It's a general reliance that, whatever the situation, he'll do the right thing: that he's trustworthy. We trust the friend's intentions, and know that his actions will be informed by those intentions.²

The second is less intimate, what sociologist Susan Shapiro calls *impersonal trust*. When we don't know someone, we don't know enough about her, or her underlying motivations, to trust her based on character alone. But we can trust her future actions.³ We can trust that she won't run red lights, or steal from us, or cheat on tests. We don't know if she has a secret desire to run red lights or take our money, and we really don't care if she does. Rather, we know that she is likely to follow most social norms of acceptable behavior because the consequences of breaking these norms are high. You can think of this kind of trust—that people will behave in a trustworthy manner even if they are not inherently trustworthy—more as confidence, and the corresponding trustworthiness as compliance.⁴

In another sense, we're reducing trust to consistency or predictability. Of course, someone who is consistent isn't necessarily trustworthy. If someone is a habitual thief, I don't trust him. But I do believe (and, in another sense of the word, trust) that he will try to steal from me. I'm less interested in that aspect of trust, and more in the positive aspects. In *The Naked Corporation*, business strategist Don Tapscott described trust, at least in business, as the expectation that the other party will be honest, considerate, accountable, and transparent. When two people are consistent in this way, we call them cooperative.

In today's complex society, we often trust systems more than people. It's not so much that I trusted the plumber at my door as that I trusted the systems that produced him and protect me. I trusted the recommendation from my insurance company, the legal system that would protect me if he did rob my house, whatever the educational system is that produces and whatever insurance system bonds skilled plumbers, and—most of all—the general societal systems that inform how we all treat each other in society. Similarly, I trusted the banking system, the corporate system, the system of police, the system of traffic laws, and the system of social norms that govern most behaviors.⁵

This book is about trust more in terms of groups than individuals. I'm not really concerned about how specific people come to trust other specific people. I don't care if my plumber trusts me enough to take my check, or if I trust that driver over there enough to cross the street at the stop sign. I'm concerned with the general level of impersonal trust in society. Francis Fukuyama's definition nicely captures the term as I want to use it: "Trust is the expectation that arises within a community of regular, honest, and cooperative behavior, based on commonly shared norms, on the part of other members of that community."

Sociologist Barbara Misztal identified three critical functions performed by trust: 1) it makes social life more predictable, 2) it creates a sense of community, and 3) it makes it easier for people to work together. In some ways, trust in society works like oxygen in the atmosphere. The more customers trust merchants, the easier commerce is. The more drivers trust other drivers, the smoother traffic flows. Trust gives people the confidence to deal with strangers: because they know that the strangers are likely to behave honestly, cooperatively, fairly, and sometimes even altruistically. The more trust is in the air, the healthier society is and the more it can thrive. Conversely, the less trust is in the air, the sicker society is and the more it has to contract. And if the amount of trust gets too low, society withers and dies. A recent example of a systemic breakdown in trust occurred in the Soviet Union under Stalin.

I'm necessarily simplifying here. Trust is relative, fluid, and multidimensional. I trust Alice to return a \$10 loan but not a \$10,000 loan, Bob to return a \$10,000 loan but not to babysit an infant, Carol to babysit but not with my house key, Dave with my house key but not my intimate secrets, and Ellen with my intimate secrets but not to return a \$10 loan. I trust Frank if a friend vouches for him, a taxi driver as long as he's displaying his license, and Gail as long as she hasn't been drinking. I don't trust anyone at all with my computer password. I trust my brakes to stop the car, ATM machines to dispense money from my account, and Angie's List to recommend a qualified plumber—even though

I have no idea who designed, built, or maintained those systems. Or even who Angie is. In the language of this book, we all need to trust each other to follow the behavioral norms of our group.

Many other books talk about the value of trust to society. This book explains how society establishes and maintains that trust.⁶ Specifically, it explains how society enforces, evokes, elicits, compels, encourages—I'll use the term *induces*—trustworthiness, or at least compliance, through systems of what I call *societal pressures*, similar to sociology's social controls: coercive mechanisms that induce people to cooperate, act in the group interest, and follow group norms. Like physical pressures, they don't work in all cases on all people. But again, whether the pressures work against a particular person is less important than whether they keep the scope of defection to a manageable level across society as a whole.

A manageable level, but not too low a level. Compliance isn't always good, and defection isn't always bad. Sometimes the group norm doesn't deserve to be followed, and certain kinds of progress and innovation require violating trust. In a police state, everybody is compliant but no one trusts anybody. A too-compliant society is a stagnant society, and defection contains the seeds of social change.

This book is also about security. Security is a type of a societal pressure in that it induces cooperation, but it's different from the others. It is the only pressure that can act as a physical constraint on behavior regardless of how trustworthy people are. And it is the only pressure that individuals can implement by themselves. In many ways, it obviates the need for intimate trust. In another way, it is how we ultimately induce compliance and, by extension, trust.

It is essential that we learn to think smartly about trust. Philosopher Sissela Bok wrote: "Whatever matters to human beings, trust is the atmosphere in which it thrives." People, communities, corporations, markets, politics: everything. If we can figure out the optimal societal pressures to induce cooperation, we can reduce murder, terrorism, bank fraud, industrial pollution, and all the rest.

If we get pressures wrong, the murder rate skyrockets, terrorists run amok, employees routinely embezzle from their employers, and corporations lie and cheat at every turn. In extreme cases, an untrusting society breaks down. If we get them wrong in the other direction, no one speaks out about institutional injustice, no one deviates from established corporate procedure, and no one popularizes new inventions that disrupt the status quo—an oppressed society stagnates. The very fact that the most extreme failures rarely happen in the modern industrial world is proof that we've largely gotten societal pressures right. The failures that we've had show we have a lot further to go.

Also, as we'll see, evolution has left us with intuitions about trust better suited to life as a savannah-dwelling primate than as a modern human in a global high-tech society. That flawed intuition is vulnerable to exploitation by companies, con men, politicians, and crooks. The *only* defense is a rational understanding of what trust in society is, how it works, and why it succeeds or fails.



This book is divided into four parts. In Part I, I'll explore the background sciences of the book. Several fields of research—some closely related—will help us understand these topics: experimental psychology, evolutionary psychology, sociology, economics, behavioral economics, evolutionary biology, neuroscience, game theory, systems dynamics, anthropology, archaeology, history, political science, law, philosophy, theology, cognitive science, and computer security.

All these fields have something to teach us about trust and security.⁷ There's a lot here, and delving into any of these areas of research could easily fill several books. This book attempts to gather and synthesize decades, and sometimes centuries, of thinking, research, and experimentation from a broad swath of academic disciplines. It will, by necessity, be largely a cursory overview; often, the hardest part was figuring out what *not* to include. My goal is to show where the broad arcs of research are pointing, rather than explain the details—though they're fascinating—of any individual piece of research.⁸

In the last chapter of Part I, I will introduce societal dilemmas. I'll explain a thought experiment called the Prisoner's Dilemma, and its generalization to societal dilemmas. Societal dilemmas describe the situations that require intra-group trust, and therefore use societal pressures to ensure cooperation: they're the central paradigm of my model. Societal dilemmas illustrate how society keeps defectors from taking advantage, taking over, and completely ruining society for everyone. It illustrates how society ensures that its members forsake their own interests when they run counter to society's interest. Societal dilemmas have many names in the literature: collective action problem, Tragedy of the Commons, free-rider problem, arms race. We'll use them all.

Part II fully develops my model. Trust is essential for society to function, and societal pressures are how we achieve it. There are four basic categories of societal pressure that can induce cooperation in societal dilemmas:

- *Moral pressure.* A lot of societal pressure comes from inside our own heads. Most of us don't steal, and it's not because there are armed guards

and alarms protecting piles of stuff. We don't steal because we believe it's wrong, or we'll feel guilty if we do, or we want to follow the rules.

- *Reputational pressure.* A wholly different, and much stronger, type of pressure comes from how others respond to our actions. Reputational pressure can be very powerful; both individuals and organizations feel a lot of pressure to follow the group norms because they don't want a bad reputation.
- *Institutional pressure.* Institutions have rules and laws. These are norms that are codified, and whose enactment and enforcement is generally delegated. Institutional pressure induces people to behave according to the group norm by imposing sanctions on those who don't, and occasionally by rewarding those who do.
- *Security systems.* Security systems are another form of societal pressure. This includes any security mechanism designed to induce cooperation, prevent defection, induce trust, and compel compliance. It includes things that work to prevent defectors, like door locks and tall fences; things that interdict defectors, like alarm systems and guards; things that only work after the fact, like forensic and audit systems; and mitigation systems that help the victim recover faster and care less that the defection occurred.

Part III applies the model to the more complex dilemmas that arise in the real world. First I'll look at the full complexity of competing interests. It's not just group interest versus self-interest; people have a variety of competing interests. Also, while it's easy to look at societal dilemmas as isolated decisions, it's common for people to have conflicts of interest: multiple group interests and multiple societal dilemmas are generally operating at any one time. And the effectiveness of societal pressures often depends on why someone is considering defecting.

Then, I'll look at groups as actors in societal dilemmas: organizations in general, corporations, and then institutions. Groups have different competing interests, and societal pressures work differently when applied to them. This is an important complication, especially in the modern world of complex corporations and government agencies. Institutions are also different. In today's world, it's rare that we implement societal pressures directly. More often, we delegate someone to do it for us. For example, we delegate our elected officials to pass laws, and they delegate some government agency to implement those laws.

In Part IV, I'll talk about the different ways societal pressures fail. I'll look at how changes in technology affect societal pressures, particularly security. Then I'll look at the particular characteristics of today's society—the Information Society—and explain why that changes societal pressures. I'll sketch what the future of societal pressures is likely to be, and close with the social consequences of too much societal pressure.

This book represents my attempt to develop a full-fledged theory of coercion and how it enables compliance and trust within groups. My goal is to suggest some new questions and provide a new framework for analysis. I offer new perspectives, and a broader spectrum of what's possible. Perspectives frame thinking, and sometimes asking new questions is the catalyst to greater understanding. It's my hope that this book can give people an illuminating new framework with which to help understand how the world works.



Before we start, I need to define my terms. We talk about trust and security all the time, and the words we use tend to be overloaded with meaning. We're going to have to be more precise...and temporarily suspend our emotional responses to what otherwise might seem like loaded, value-laden, even disparaging, words.

The word *society*, as used in this book, isn't limited to traditional societies, but is any group of people with a loose common interest. It applies to societies of circumstance, like a neighborhood, a country, everyone on a particular bus, or an ethnicity or social class. It applies to societies of choice, like a group of friends, any membership organization, or a professional society. It applies to societies that are some of each: a religion, a criminal gang, or all employees of a corporation. It applies to societies of all sizes, from a family to the entire planet. All of humanity is a society, and everyone is a member of multiple societies. Some are based on birth, and some are freely chosen. Some we can join, and to some we must be invited. Some may be good, some may be bad—terrorist organizations, criminal gangs, a political party you don't agree with—and most are somewhere in between. For our purposes, a society is just a group of interacting *actors* organized around a common attribute.

I said actors, not people. Most societies are made up of people, but sometimes they're made up of groups of people. All the countries on the planet are a society. All corporations in a particular industry are a society. We're going to be talking about both societies of individuals and societies of groups.

Societies have a collection of *group interests*. These are the goals, or directions, of the society. They're decided by the society in some way: perhaps formally—either democratically or autocratically—perhaps informally by the group. International trade can be in the group interest. So can sharing food, obeying traffic laws, and keeping slaves (assuming those slaves are not considered to be part of the group). Corporations, families, communities, and terrorist groups all have their own group interests. Each of these group interests corresponds to one or more norms, which is what each member of that society is supposed to do. For example, it is in the group interest that everyone respect everyone else's property rights. Therefore, the group norm is not to steal (at least, not from other members of the group⁹).

Every person in a society potentially has one or more *competing interests* that conflict with the group interest, and *competing norms* that conflict with the group norm. Someone in that we-don't-steal society might really want to steal. He might be starving, and need to steal food to survive. He just might want other people's stuff. These are examples of *self-interest*. He might have some competing *relational interest*. He might be a member of a criminal gang, and need to steal to prove his loyalty to the group; here, the competing interest might be the group interest of another group. Or he might want to steal for some higher moral reason: a competing *moral interest*—the Robin Hood archetype, for example.

A *societal dilemma* is the choice every actor has to make between group interest and his or her competing interests. It's the choice we make when we decide whether or not to follow the group norm. Those who do *cooperate*, and those who do not *defect*. Those are both loaded terms, but I mean them to refer only to the action as a result of the dilemma.

Defectors—the liars and outliers of the book's title—are the people within a group who don't go along with the norms of that group. The term isn't defined according to any absolute morals, but instead in opposition to whatever the group interest and the group norm is. Defectors steal in a society that has declared that stealing is wrong, but they also help slaves escape in a society where tolerating slavery is the norm. Defectors change as society changes; defection is in the eye of the beholder. Or, more specifically, it is in the eyes of everyone else. Someone who was a defector under the former East German government was no longer in that group after the fall of the Berlin Wall. But those who followed the societal norms of East Germany, like the Stasi, were—all of a sudden—viewed as defectors within the new united Germany.

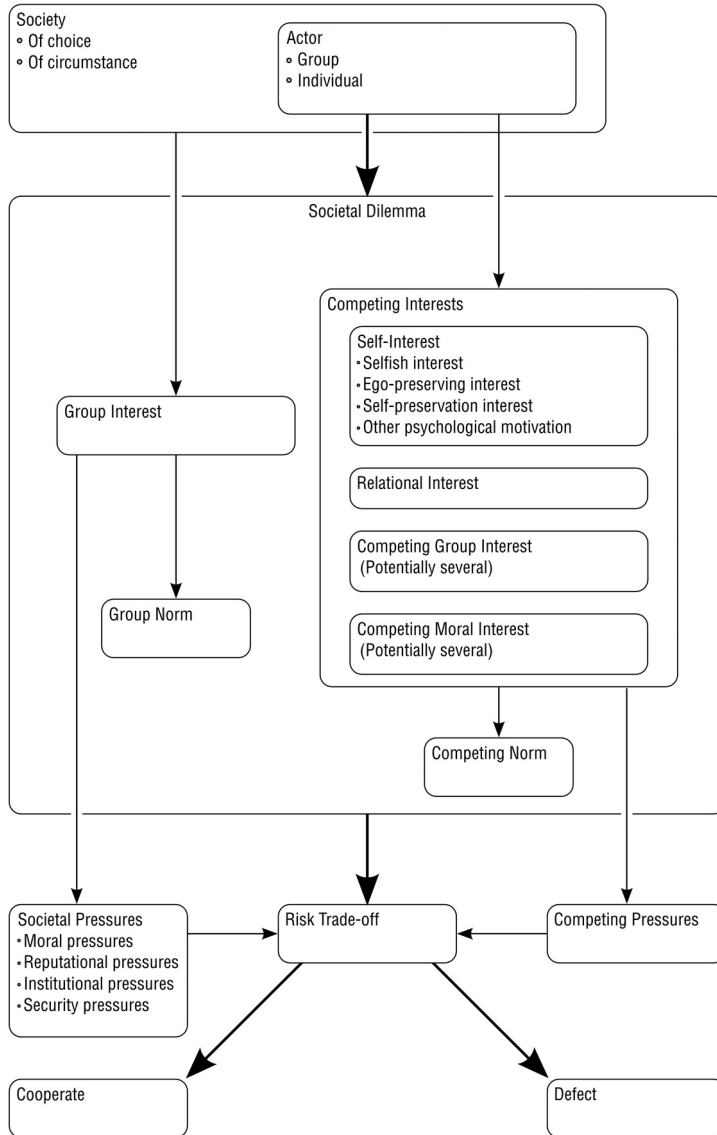


Figure 1: The Terms Used in the Book, and Their Relationships

Criminals are defectors, obviously, but that answer is too facile. Everyone defects at least some of the time. It's both dynamic and situational. People can cooperate about some things and defect about others. People can cooperate with one group they're in and defect from another. People can cooperate today and defect tomorrow, or cooperate when they're thinking clearly and defect when

they're reacting in a panic. People can cooperate when their needs are cared for, and defect when they're starving.

When four black North Carolina college students staged a sit-in at a whites-only lunch counter inside a Woolworth's five-and-dime store in Greensboro, in 1960, they were criminals. So are women who drive cars in Saudi Arabia. Or homosexuals in Iran. Or the 2011 protesters in Egypt, who sought to end their country's political regime. Conversely, child brides in Pakistan are not criminalized and neither are their parents, even though in some cases they marry off five-year-old girls. The Nicaraguan rebels who fought the Sandinistas were criminals, terrorists, insurgents, or freedom fighters, depending on which side you supported and how you viewed the conflict. Pot smokers and dealers in the U.S. are officially criminals, but in the Netherlands those offenses are ignored by the police. Those who share copyrighted movies and music are breaking the law, even if they have moral justifications for their actions.

Defecting doesn't necessarily mean breaking government-imposed laws. An orthodox Jew who eats a ham and cheese sandwich is violating the rules of his religion. A Mafioso who snitches on his colleagues is violating *omertà*, the code of silence. A relief worker who indulges in a long, hot shower after a tiring journey, and thereby depletes an entire village's hot water supply, unwittingly puts his own self-interest ahead of the interest of the people he intends to help.

What we're concerned with is the overall *scope of defection*. I mean this term to be general, comprising the number of defectors, the rate of their defection, the frequency of their defection, and the intensity (the amount of damage) of their defection. Just as we're interested in the general level of trust within the group, we're interested in the general scope of defection within the group.

Societal pressures are how society ensures that people follow the group norms, as opposed to some competing norms. The term is meant to encompass everything society does to protect itself: both from fellow members of society, and non-societal members who live within and amongst the society. More generally, it's how society enforces intra-group trust.

The terms *attacker* and *defender* are pretty obvious. The predator is the attacker, the prey is the defender. It's all intertwined, and sometimes these terms can get a bit muddy. Watch a martial arts match, and you'll see each person defending against his opponent's attacks while at the same time hoping his own attacks get around his opponent's defenses. In war, both sides attack and defend at the tactical level, even though one side might be attacking and the other defending at the political level. These terms are value-neutral. Attackers can be criminals trying to break into a home, superheroes raiding a criminal

mastermind's stronghold, or cancer cells metastasizing their way through a hapless human host. Defenders can be a family protecting its home from invasion, the criminal mastermind protecting his lair from the superheroes, or a posse of leukocytes engulfing opportunistic pathogens they encounter.

These definitions are important to remember as you read this book. It's easy for us to bring our own emotional baggage into discussions about security, but most of the time we're just trying to understand the underlying mechanisms at play, and those mechanisms are the same, regardless of the underlying moral context.

Sometimes we need the dispassionate lens of history to judge famous defectors like Oliver North, Oskar Schindler, and Vladimir Lenin.