Chapter 3: The Prisoner's Dilemma and the Emergence of Cooperation

ABSTRACT: John Nash's concept of equilibrium demonstrates how games may have sub-optimal solutions that are nevertheless stable, because neither player can improve their condition unilaterally. Advertising is inherently a sub-optimal condition for both players: marketers would prefer to win consumers without spending money on advertising, and consumers would prefer to enjoy content without being advertised to. The concept of the Prisoner's Dilemma illuminates this condition, because it suggests that consumers and marketers could reach a more satisfying relationship if they could coordinate cooperation. Studies conducted on iterative rounds of the Prisoner's Dilemma demonstrate this natural evolution toward cooperation and reveal a set of conditions that make cooperation possible, but they also demonstrate the fragility of cooperation and the potential for downward spirals of mutual defection.

Just as banner advertising could not evolve into a more nuanced medium until its near-death experience forced it to do so, marketing in general does not willingly forego its short-term gains in favor of long-term stability. Put simply, restraint does not occur in the absence of consequence, and so the advertiser will push their advantage in reaching a set of consumers until that negative consequence is achieved in the form of diminished returns.

This tendency may be dramatized as greed or blindness in the face of consumer resentment, but in fact it is perfectly logical and explainable within the terms offered by game theory. As noted before, consumers would find it optimal to find good products and enjoy free media content without being marketed to at all, and marketers would find it optimal for consumers to choose the marketer's products over others without the need to spend a single dollar on advertising. But both sides compromise in the interest of achieving their goals and find a point of equilibrium. In this important respect, *advertising is always the pursuit of the sub-optimal*, a means to an

end for both players. For the consumer it is a Faustian bargain, which puts the marketer in the unfortunate role of Mephistopheles.

The natural tendency of each side to push their respective advantage is illustrated in the traditional "S" curve, by which the effectiveness of a given media spend is often evaluated and optimized. The S curve indicates the impact on sales or some other success metric (the "Y" axis) of a given number of media impressions (the "X" axis). The upper arc of the curve indicates the point of equilibrium – the optimal number of impressions necessary to achieve the best possible sales outcome. The effectiveness of further impressions is diminished beyond that point.
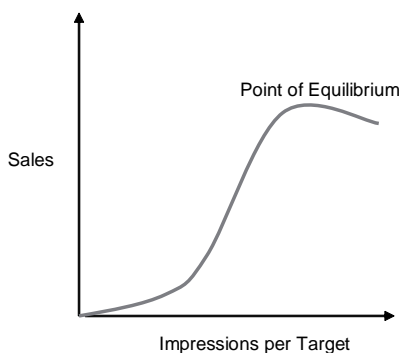

Figure 3: Marketing efficiency "S" curve

Thus the marketer has a built-in incentive to pursue the maximum number of impressions possible, provided they produce an incremental return.

In greatly simplified terms, the marketing industry as a whole operates as one enormous "S" curve, pursuing an advantage to its furthest logical point. And since that industry doesn't operate as a single entity but rather as a vast array of independent players, there is no collective incentive to change course or to sacrifice short-term gains for long-term health.

There are, in fact, specific conditions under which a group of independent players in a game will make such sacrifices; these will be explored in Chapter Five's examination of the coordination game. For now I will stipulate that these conditions do not presently exist in marketing to any widespread degree. If they did, marketers would not find themselves trapped in a dialectic of hype and backlash. This dialectic occurs because of the instability in the marketer-consumer equilibrium, in which marketers continually press their advantages to compensate for consumers' increasing disdain for their messages.

This equilibrium has always been delicate at best; it means that advertising impressions and response are in a symbiotic balance, with consumers

tolerating enough advertising to grant them access to free or subsidized content, and marketers are gaining sufficient return from this advertising to limit the number of impressions they impose on consumers. Upsetting the equilibrium would cost the marketer more money and expose the consumer to more advertising – a less optimal outcome for both. Neither has an incentive to defect unilaterally.

3.1 The Great Consumer Opt-Out

But like all equilibria in game theory, this one is upset by shifts in each player's knowledge and opportunity. Fully analyzing these shifts would require deeper historicizing of the chicken-and-egg relationship between marketing saturation and consumer disdain than this book will attempt, but suffice it to say that marketers have, in the last two decades, gained vastly greater access to advertising opportunities than in previous eras, and this has upset the balance. We often hear that consumers in the U.S. are exposed to more than 3,000 ad messages per day (Taylor), but the real number, accounting for logo and label exposure, product placement, etc., is probably several times higher. Many of these opportunities are newly minted: not only Web advertising, but commercial email, naming rights, product placement, mobile advertising, and on and on.

The result of this act of defection is a corresponding defection on the part of consumers, with increasing ad tune-out and opt-out. The explosion in non-traditional means of advertising like product placement is, in part, a direct result of marketers seeking alternatives to television advertising, which has been severely compromised by the growth of Tivo and other ad-skipping technologies. One study showed that 90% of consumers that can skip television ads do so, an act of defection made possible by a shift in opportunity – a technological one, in this case – in the consumer's favor (Pasik). And so we find ourselves in a downward spiral of defection, with ad exposure and ad tune-out accelerating at a corresponding rate.

The degree of consumer inurnment to ad messages alone is sufficient to demand a shift in strategy for marketers away from the zero-sum game. To cite a few examples: A study by the Stanford Poynter Institute in 2000, a mere six years after banner advertising's debut, uncovered the phenomenon of "banner blindness," whereby Web users develop the ability to tune out advertising on Web pages they were viewing. Participants in the study saw banner ads only 45% of the time, with an average attention of only 1 second. Since banner ads typically take several seconds to deliver a message, investing in this level of attention is a bad deal for marketers, to say the least.

This tendency has only worsened over time: Web usability expert Jakob Nielsen has conducted multiple banner blindness studies since 2000, using heat-tracking technology to record users' eye movements, and has reached the grim conclusion that "Users almost never look at anything that looks like an advertisement, whether or not it's actually an ad" (Nielsen).

Email is similarly besieged by indifference. Arguably this has much to do with the increase in illegal spam email, which cannot be attributed to bad behavior by legitimate marketers, but the net impact is the same. The anti-spam company Postini reported in 2008 that 94% of all email was spam, with the rate of spam increasing by 1.2% *per day* (Keizer). This has contributed to the discrediting of legitimate emailers, with only 20% of consumers saying they trust email that they've opted in to receive.

In this downward spiral, both players are attempting to regain their position by unilateral defection: the marketer by increasing the number of impressions or ad exposures, and the consumer by decreasing attention. This is a poor strategy for both. The marketer's credibility further erodes, and the consumer is merely advertised to in more and increasingly pernicious ways in order to compensate for the loss of attention. It's also worth pointing out that the consumer's defection also harms their ability to enjoy sponsored content, in a variety of meaningful ways. Some examples:

- Newspapers facing declining revenues reduce their national and global coverage and cut back investigative reporting, thus producing less useful content
- Television programming increasingly relies on cheaper formats like reality programs, reducing the variety and quality of content
- Programming across channels that fails to attract an immediate audience is shelved more quickly, further reducing content variety

When mapped to a payoff matrix, this mutually unsatisfying arrangement becomes obvious (Table 4). The downward spiral is represented by the upper right quadrant, with the marketer increasing their spending, the consumer ignoring the messaging, and the content suffering for it. For the consumer, responding to the ads is marginally better, because doing so increases the health of sponsored content and may prompt cooperation – in the form of decreased exposure – from the marketer. Correspondingly, the optimal payoff for the marketer involves spending less on marketing but getting greater response.

Table 4: Advertising response payoff matrix

	Consumer: *Respond to advertising*	*Consumer:* *Ignore advertising*
Marketer: Increase exposure	3-*2*	1-*1*
Marketer: Decrease exposure	4-*4*	2-*3*

It scarcely requires pointing out that such behavior by both parties – consumers voluntarily responding to more ads and marketers voluntarily decreasing spending in order to get on better terms with each other – is patently absurd. But that is precisely why this game has two points of equilibrium – the worst case scenario in the upper right, and the best case in the lower left. Recall that a point of equilibrium is not defined as the optimal solution; it is merely a point at which neither player can improve their position by acting unilaterally. In the downward spiral, both players are defecting, but unilateral cooperation accomplishes nothing. It is absurd for the marketer to think that lower ad exposure will unilaterally produce a higher response, and it is equally absurd for the consumer to think that unilaterally responding to more ads will ease the ad bombardment or improve content quality. And so the downward spiral continues.

But the game has another point of equilibrium. In the lower left quadrant, the marketer and the consumer find symbiosis. The marketer puts fewer, more relevant messages in front of the consumer, and the consumer responds more frequently. Content quality improves without the need for heavier sponsorship. Neither side has an incentive to defect unilaterally, because all such moves produce sub-optimal solutions.

Getting to this point requires a degree of cooperation that can't occur in the absence of outside factors. And this is precisely where social media enters the picture as the outside factor with the power to change the game. While it has become fashionable to speak of social media as a fast-rising groundswell that began to take shape around 2007, it is more accurate to regard it as the culmination of forces that have been endemic to the Web from the beginning. For consumers, its early manifestations were in tools like product ratings & reviews, message boards, chats, and newsgroups, all of which had been in place for more than a decade. These foundational elements were in fact essential to social media's more recent explosion as a full-blown cultural phenomenon, because they conditioned consumer behavior and expectations to recognize that greater participation and transparency were available to them than ever before.

3.2 The Shifting Ground of Consumer Trust

The key element in social media is engagement, and to the same extent that consumers were disengaging in advertising in advertising, they were learning to engage each other. The global PR firm Edelman has been conducting an annual "Trust Barometer" survey to gauge consumer trust in various information sources for the last 9 years, and their 2006 report offered a startling finding: Trust in "a person like me" had risen from only 20% in 2003 to 68%, surpassing all other sources, including doctors and academics. The company's CEO, Richard Edelman, issued the prophetic admonition that "companies need to move away from sole reliance on top-down messages delivered to elites toward fostering peer-to-peer dialogue among consumers and employees, activating a company's most credible advocates" (Edelman).

How could something as fundamental as trust shift so dramatically in only 3 years? Loss of trust in traditional media sources was certainly a contributor; the same Barometer report noted that trust in television as a first source of trustworthy information had fallen from 39% to 29% in two years. But this loss does not automatically privilege trust in "a person like me." Television, after all, is ubiquitous, and like-minded peers with expertise on specific subjects are hard to come by. Or at least they use to be.

Significantly, the same report noted that 34% of consumers take action against a distrusted company by sharing "negative company opinions/experiences online." Therein lies the sea change in behavior that begat a sea change in trust. Consumers gradually, and now naturally, gained easy access to the means of providing feedback online, and their fellow consumers grew to trust these new media as information sources.

Suppose, for instance, that I am planning a trip to Honolulu back in 1996, and I want to find a good beachfront resort. I can rely on travel guides, which will offer me only one perspective on a given resort. I can look at travel magazines, which may contain content about the resorts – sponsored by the same resorts. And I can obtain brochures from the resorts themselves. It's unlikely that I'd be able to gather enough knowledge about a given resort – or even about Honolulu – from friends and family to be able to count on this peer group as an information source at all.

But a decade later, all of these sources are vastly overshadowed – if they are even consulted at all – by a single Web site: TripAdvisor.com, which can offer me dozens of meticulous, detailed opinions and ratings on each resort in Honolulu. I also have access to Honolulu message boards containing peer advisors with local expertise. These opinions are unsolicited, unpaid for, and frank, and most significantly, *there are 20 million of them*. In a keystroke,

my trust and reliance on less available, less thorough, and less objective sources plummets.

My point is not that TripAdvisor offers something unique in its subversion of the traditional means of information-gathering about a potential purchase; indeed, it is only the most prominent of many examples of travel review sites. It is the subversion itself that is the game-changer. We can all agree without much need for further analysis that unpaid, unsolicited first-hand accounts of a destination or product are going to be more reliable and therefore more valuable to the consumer than advertiser-supplied information. A single opinion might be too anecdotal to trust, but a dozen opinions have serious weight. So it follows that the ubiquitous availability of such information changes the advertising game in two important ways:

1. In zero-sum games, it provides an informational advantage to the consumer that shifts the equilibrium point in their favor, e.g., on product pricing and discounts, because it provides them with knowledge of the marketer's moves.
2. In non-zero-sum games, which will be the subject of this chapter, consumers' traditional reliance on marketers for product information has been vastly reduced or eliminated, so advertising itself is further diminished in value.

The importance of this latter point cannot be overstated. The essence of the game is that the consumer and marketer are *mutually dependent* adversaries, but the removal of one key area of dependence – if the consumer wanted to learn about the product, they had to hear from advertisers – has radically shifted the game in favor of the consumer.

3.3 The Marketer's Loss of the Informational Advantage

To understand the implications of this, let's return to the example of TripAdvisor and contemplate the dilemma from the point of view of a Honolulu hotel marketer facing a spate of mixed or poor reviews.

To begin with, as the marker I cannot counter this negative information simply by choosing a different playing field, i.e., by reaching the consumer by other means. It is nearly inevitable that the interested consumer will at some point take to the Web to learn more about my hotel. If I'm very concerned about my reviews, I can counter-weight this tendency by focusing on consumers that are less likely to read online reviews – senior citizens, for instance. (Music labels have followed a similar strategy in promoting artists

whose demographic appeal makes their music less likely to be illegally downloaded.) But trading a wide audience for a narrow one still constitutes a shift in equilibrium to the marketer's disadvantage.

By the same token, I cannot simply outspend the problem, because I'm not operating on a playing field that makes that option viable. In the so-called "Cola Wars" of the 1980's, Pepsi-Cola and Coca-Cola played the game in the advertising arena and maintained an equilibrium, in effect, as two marketers posting entirely biased but entertaining reviews of each other's products; the consumer's role was secondary and passive. In this new paradigm, the consumer's role is primary. If my competitors achieve better consumer reviews, then overspending on marketing will not solve the problem, and I am potentially made *more* vulnerable by expending resources in a futile effort.

I could also try to counteract the results of negative reviews by discounting. If my hotel is overpriced at $200 a night, perhaps it is a bargain at $100 a night. But this obviously constitutes a shift in equilibrium by the zero-sum terms defined in the last chapter, and it does nothing to counteract the negativity directly. Plus I lose money.

Even if the consumer does a direct search for my hotel, outside opinions are inescapable. (Try this exercise with any hotel. A search for "Ritz Carlton Cancun" on Google returns the hotel's Web site as the top result. The #2 result? The Ritz Carlton Cancun review page on TripAdvisor). I might try advertising on the TripAdvisor site, and if my reviews are good, this would be a highly effective strategy; consumers could go directly from interest to action by reading the reviews then clicking the ad to visit my site. But if the reviews are uniformly good, I never had much to worry about in the first place, and the fundamental alteration remains the same: I have no course of action available to me in traditional marketing to counter the consumer's information-gathering move.

The conclusion is inescapable: as a hotel marketer I am going to have to move outside of my traditional marketing comfort zone and take on the issue of negative reviews head-on. This is going to become as integral to my hotel's marketing strategy as my logo, and it isn't going to be easy. I will work with customer service to respond to negative reviews and elicit positive testimonials. I will coordinate customer care initiatives to address the issues that led to bad reviews in the first place. I will send emails and mailers to past guests soliciting positive responses and surveying their experience. I will tirelessly monitor, catalogue, and respond to customer reactions to my hotel, and I will track the uptick in positivity.

It is altogether obvious that I am describing a very different kind of marketing activity than the ones marketers are accustomed to, and certainly

different from zero-sum. The specific tactics used in this new marketing game will be explored in far greater detail in later chapters; my primary purpose in offering the hotelier example here is to underscore the previously-described limitations of zero-sum and set the stage for a different game theory concept that will help us analyze the challenges faced by the hotelier, and indeed, all marketers in this new era.

As noted, the basic limitation of zero-sum is that presupposes one player's direct gain is the other player's direct loss, and even traditional marketing doesn't usually work that way. Marketers and consumers are mutually dependent because consumers want products and marketers want to sell them, and advertising the right product to the right consumer at the right time constitutes a mutual gain for both players. But as has also been previously described, this entire arrangement has an equilibrium solution that is *always sub-optimal*, always compromised. The optimal solution for marketers would be to sell their products over other products with no investment in advertising, and the optimal solution for the consumer would be to select the right product over other products (and to be able to enjoy free media content) without being exposed to advertising. Neither of these optimal outcomes has long-term viability, and so consumers and marketers do their endless dance.

3.4 The Prisoner's Dilemma

The most famous of all game theory concepts, the Prisoner's Dilemma, is precisely what this situation demands, because it is most often used to analyze the conditions of cooperation and defection in situations involving sub-optimal solutions. As a "dilemma," it is much more complex and nuanced than zero-sum, and so it is worth exploring at some length.

The Prisoner's Dilemma has its origins in a paper produced by a pair of game theory's original practitioners, Merrill Flood and Marvin Dresher, in 1949. The paper offered a set of real-life scenarios that were meant to explore the limits of the Nash equilibrium (Dixit & Nalebuff 2008). You'll recall that the Nash equilibrium posits that for every two-person game, there is at least one stable equilibrium point in which *neither player can improve their outcome unilaterally*, given the moves available to the other player. As we saw in the example of Manufacturer X, the point of equilibrium can shift if one player gains information about the other's available moves, but those *information gains always favor one player over the other*. If the consumer learns that Manufacturer X is willing to sell the product for $200 less than they were prepared to pay, that insight helps the consumer and hurts the

manufacturer. For the traditional hotel advertiser facing poor reviews on TripAdvisor, the availability of those reviews hurts the advertiser and helps the consumer.

The unique feature of the Prisoner's Dilemma is that *both* players accept a sub-optimal solution based on the information they have, but *both* could have arrived at an optimal solution had they been able to coordinate their moves. This potentially turns Nash's equilibrium on its ear, because it uncovers a basic instability: it suggests that in some scenarios, coordinating information-sharing could produce better outcomes for both players. Doing so can be tricky, but the rewards may be worth it.

To understand this, let's look at the dilemma itself. Examples abound, but the most famous is the dilemma's namesake, articulated by game theorist Albert Tucker. Imagine two criminal conspirators are arrested for a robbery. The two are separated by the police, and in a set-up familiar to anyone who has seen a prime-time police drama, each prisoner is invited to implicate the other in order to receive a lighter sentence (Tucker 1983).

The game presupposes that the police need one of the players to implicate the other in order to get a conviction on all charges; if both players stay silent, both players will receive only a one-year sentence for lesser charges. But the players are separated; they cannot coordinate their actions. If one player implicates the other while the other stays silent, the silent one will get a harsh sentence – 6 years – while the other players goes free. If both players implicate the other, no ringleader is established, and both receive lighter sentences of 3 years each.

These stark options are easy to articulate in a payoff matrix if we reverse the polarity of the numbers so they refer to the years in a prison sentence; a zero means no prison sentence, and so on.

Table 5: Prisoner's Dilemma payoff matrix

	Prisoner 2: Stays silent	Prisoner 2: Betrays
Prisoner 1: Stay silent	1-*1*	6-*0*
Prisoner 1: Betrays	0-*6*	3-*3*

According to the payoff matrix, both players are better off defecting (betraying), no matter what the other player does, because they have no way of coordinating what the other player will do. Betraying will result in either no prison time or three years, whereas staying silent carries the risk of 6 years behind bars. In this case, defecting is a dominant strategy for both players

and produces a Nash equilibrium in the lower right quadrant (3-3), because neither player can unilaterally improve on this position.

But the dilemma is a true dilemma for several reasons. First, information-sharing (coordination) would produce a better solution for both players (the 1-1 outcome in the upper left quadrant), which runs contrary to Nash's theorem. If only the prisoner's could signal each other, their lot would be *vastly* improved!

Secondly, and more importantly for our purposes, when this dilemma is taken out of the laboratory and viewed through the lens of human emotion rather than pure logic, the sub-optimal solution is extremely unsatisfying. Game theory teaches us to "Hope for the best, but prepare for the worst" (Dixit & Nalebuff 2008), but in real-life situations, hope often seems to trump preparation. If there is any honor at all among thieves, their sense of morality – typically excluded from game theory analysis – would rebel against betraying their partner in crime. And then there is the starkness of it all – a very light one-year sentence is such a vast improvement over a 3-year sentence that one can hardly bear to imagine that one's fellow thief would fail to reach the same conclusion and unilaterally cooperate.

It should be no great surprise, then, that reckless pursuit of the optimal instead of the more stable sub-optimal is often what we see when the Prisoner's Dilemma is played outside of the lab. In the real-world environment, we find two different PD scenarios at work, each demanding different strategies: one-off games and the more common and important "Iterated Prisoner's Dilemma."

3.5 The Hidden Allure of Cooperation

Pure logic and conventional wisdom dictate that in one-off games of Prisoner's Dilemma – as in the example above – *it is always better to defect*. That is the dominant strategy. It follows, then that we would see this result borne out in the countless versions of PD staged by academics over the decades.

But in fact we see the opposite. Dixit & Nalebuff report in that in aggregate across one-off PD experiments, cooperation occurs almost half the time, "even when each pair of players meets only once" (Dixit & Nalebuff 2008). They offer the fascinating example of the TV game show, *Friend or Foe*, in which players competing for a pot of money were simply required to secretly write down "Friend" or "Foe" to indicate their move. If one player cooperated and the other defected, the defector got the whole pot. If both cooperated, they split the pot. If both defected, they got nothing.

Applying the strict logic of "preparing for the worst," it is quite obvious that defecting – choosing "foe" – is the better strategy in a single round, because you'll either end up with the full pot or nothing. Choosing "friend" nets you half the pot or nothing, and of course you have no way of knowing if your opponent wishes to cooperate. Yet Dixit & Nalebuff report that almost half of the contestants chose "friend," preparing for the best instead. When I posed this dilemma to my 8-year-old son, he immediately chose "friend," and I was torn between admiration for his altruistic tendencies and chagrin at his hasty reasoning. (Attempting to iterate only seemed to annoy him). Or perhaps he simply wasn't paying attention to the game, which may be the most common scenario for marketers too.

The result is clear: we tend toward cooperation even when the odds are stacked against us. But why? The answer necessarily lies outside of an ultra-rational application of game theory, since it seems to involve either a degree of selflessness or some different transaction – hidden terms in the game, as it were. Such an answer would also help us explain why travelers visiting TripAdvisor would devote so much time – 20 million reviews and counting! – to sharing their experiences with other travelers, with no tangible reward at stake.

This psychological basis for cooperation will be explored thoroughly in the next chapter, when we dig into examples of cooperation in the social media arena. For now, let's also consider how and why cooperation emerges in iterated games of Prisoner's Dilemma. The difference is important, because while cooperation may occur in one-off games as the result of altruism, blind hope, or some other emotional cause, it is often a rational response in iterated PD.

3.6 The Iterated Prisoner's Dilemma

Iterated PD occurs when two players face each other in consecutive matches, with the outcome decided by the cumulative score. Most real-life dilemmas are in fact iterative; the U.S. and the Soviets negotiated many arms agreements, and marketers and consumers square off thousands upon thousands of times. It leads us to the question of what strategy is best *in the long run* rather than in short-term self-interest.

The interesting feature of iterated PD is that it brings in punishment or retaliation as a feature; one might reconsider the logic of defection if it makes your opponent more likely to retaliate in the next round. The theory's originators, Flood and Dresher, uncovered this insight when they staged an iterative PD game among their game theory colleagues, featuring the same

two players squaring off 100 times in a single session. The session revealed a "difficult struggle to secure mutual cooperation," (Poundstone 1992) which was logical, since both players stood to gain more if they could coordinate their actions and bring about iterative cooperation. When one player defected, the other would punish him with a defection in the next round, and both would return to cooperating thereafter.

In all, the supposed Nash equilibrium – mutual defection – occurred only 14 times in 100 rounds of play. When Flood and Dresher showed these results to John Nash, he complained that the entire set-up was more like one multi-move game, in which his theory would not apply, because "'There is too much interaction, which is obvious in the results of the experiment'" (Poundstone 1992). Indeed, but it is precisely such interaction that interests us when we apply the Prisoner's Dilemma to the marketing world, because the interaction involves an exchange of information about the other's intentions.

In the Flood and Dresher experiment, each player's direct intentions in each round was still kept hidden, but each player could glean some insight into the other's future intentions based on their past actions, in exactly the same way that poker players observe when other players have a tendency to bluff, even if their actual hand remains hidden. Thus information still changes hands, even if it is not complete information.

We see this play out in consumer responses to various forms of marketing. Consumers' wholesale defection from banner advertising in the early aughts was a direct response to the perceived defection on the part of marketers. Consumers had been subjected to ads with fake interfaces, which led to unintended clicking, as well as pop-unders, pop-overs, and all manner of dirty tricks. While the majority of advertisers did not engage in such tactics, the impact of the defection was wholesale; consumers mistrusted the intentions of online advertisers in general. They stopped responding, and it required many rounds of cooperation over many years for the relationship to regain its equilibrium.

3.7 The Persistent Problem of Bad Apples

The consequence of early defection, after which many rounds of cooperation are necessary to rebuild trust, is one of the key insights of the iterated Prisoner's Dilemma. Simply put, defection is a short-term gain but a long-term loss. It leaves a poor first impression, or at least an impression that one is a defector by nature, which is a bad thing to be in an iterated Prisoner's Dilemma. Thus Poundstone describes the primary ingredient of

the Prisoner's Dilemma as "a temptation to better one's interests in a way that would be ruinous if *everyone* did it" (Poundstone 1992).

And therein lies the rub. In marketing, successful cooperation requires not only that the marketer signal cooperation to the consumer, i.e., "If I you click on this banner ad, you can trust me that it will not trick you," but the marketer must also somehow reinforce cooperation among his fellow marketers, i.e., "We will not fool consumers, and we will punish those who do."

This is the classic problem of the bad apples. In the statistics I cited earlier regarding distrust for commercial email, it is clear that consumer disgust with the bad apples – the illegal spammers – has indeed spread to the whole bunch, and not without reason: marketers are continuously probing the gray areas left open by CAN-SPAM restrictions, once again building a reputation for defection. Similarly, consumer disgust with telemarketing phone calls – many of which were misleading and overly persistent – led to the creation of the National Do Not Call Registry in 2003. Telemarketers had been engaged in iterative acts of defection – reaching consumers in a way that they actively disliked – and consumers finally defected en masse in retaliation. An astonishing 72% of Americans had placed themselves on the list by 2007 (Federal Trade Commission 2007), easily the single greatest act of consumer revolt in the history of marketing. Legitimate marketers vehemently opposed the legislation, arguing that it amounted to punishing all for the acts of a few. And indeed it did, but such is the retributive nature of the iterated Prisoner's Dilemma.

So the tendency of marketers to shoot themselves in the foot has a strong historical basis, even if the benefits of cooperation are equally well established. It is far too easy for individual marketers to act in their own short-term interests; it is much harder, in a highly competitive landscape, to act in the long-term interests of the common good. How, then, could cooperation possibly emerge?

The simple answer is that it emerges in the context of a set of rules that resolve the Prisoner's Dilemma by ensuring that players can properly signal their intentions and cooperation can flourish. But significantly, these have to be rules within the game itself, i.e., between players, and not ones imposed from the outside. Marketers fought legislation like CAN-SPAM and the Do-Not-Call Registry because they naturally feared that such rules would overly restrict their ability to play the game. Rules that emerge in the context of consumer response work the best, because marketers have to pay attention to the moves of their opponent in order to play the game.

What marketers really need is for consumers to be able to signal their cooperative response, i.e., they need the information coordination described

earlier. If the emergent rules aid the transfer of information, so that marketers feel assured of consumer cooperation, then they themselves are more likely to cooperate. Is it possible to play the game with this level of clarity and conviction? Generally, yes. It has been accomplished in the groundbreaking iterated Prisoner's Dilemma strategy known as TIT FOR TAT.

3.8 The Enduring Relevance of TIT FOR TAT

TIT FOR TAT is the product of a competition conducted by the political science professor Robert Axelrod in 1980 and described in his 1984 work *The Evolution of Cooperation*. Axelrod invited academics from the fields of psychology, economics, political science mathematics, and sociology – all of them familiar with the Prisoner's Dilemma – to submit computer programs that would play iterated games of the Prisoner's Dilemma against each other. Each program would be pitted against another, round robin style, and 200 rounds would be played in each match-up. Because the games would be played by computer, emotional effects – like hoping for cooperation – could not taint the results.

Fourteen programs were submitted in Axelrod's first experiment. Many were highly sophisticated, comprising dozens of lines of code. TIT FOR TAT was astonishingly simple, consisting of only 4 lines of code and a strategy so basic that a kindergartner could play it: *Cooperate in the first round. After that, do whatever your opponent did in the last round.*

TIT FOR TAT not only handily won Axelrod's tournament, it won a second tournament among 62 contenders that tried to improve on its initial success, and it has never lost its first-place status after three decades. Its logic is unassailable. It offers no pattern that an opponent can exploit (recalling the importance of randomization in iterative games), because it simply responds to its opponent's actions. The opening cooperative move sets a positive agenda that encourages further cooperation, but if the opponent defects, he is continually punished until cooperation is regained.

Axelrod offers useful conclusions as to what makes TIT FOR TAT such a successful strategy. TIT FOR TAT works, Axelrod claims, because it is "nice, retaliatory, forgiving, and clear." One could hardly hope for a more useful and succinct explanation of how to succeed in mutually dependent conflicts. Axelrod goes on to show how these same features could be found in other conflicts that engendered cooperation, including the détentes that emerged in WWI's trench warfare, when enemy soldiers refused to fire on each other, and in the evolution of biological systems. The beauty of TIT FOR TAT, in Axelrod's view, is that it succeeded in engendering better

behavior in its opponents, i.e., it elicited "behavior from the other player that enabled both to do well" (Axelrod 2006).

As we'll see in the next chapter, this overall elevation in the level of discourse and behavior on both sides is critical to the success of social media marketing. We would do well to remember that these strategies are not the least bit altruistic; they serve the interests of both sides, and there is punishment awaiting the defector (social media offers ample means of punishment, as we'll see). So TIT FOR TAT is no patsy strategy, but it also hopes for the best, taking the risk of a cooperative opening move. This turns out to be key: Axelrod reports that "the single best predictor of how well a rule performed was whether or not it was nice." Niceness was a feature of all of the top eight performers, and none of the bottom seven. From that we can derive another rule that will be critical to marketers approaching the social media space: *In an iterative game, never be the first to defect.*

3.9 The Dangers of the Death Spiral

The above rule actually underscores the one potential weakness in the TIT FOR TAT strategy. It is simply this: mistakes can be fatal. In a TIT FOR TAT software program, the chances of a glitch causing an erroneous response are rather low; it is, after all, a very simple piece of logic. But if the program did manage to defect by accident in response to cooperation on the previous move, it could prompt the opponent to defect in response, resulting in another defection, and so on. This outcome has become known as the "death spiral," and it's a popular device in Hollywood movies like *Reservoir Dogs*, when the characters find themselves in an armed stand-off and some accidental stimulus causes everyone to shoot each other (Tarantino 1992). It's also a distinct possibility when marketers are playing real-life TIT FOR TAT; the history of banner advertising is something like this death spiral scenario.

I mention the death spiral because it will be important in examining social media blunders in the next chapter. It's relatively easy for well-intentioned marketers dabbling in the social media space to overstep their bounds and provoke rather vitriolic consumer backlash. In general, these consumers are not being intolerant; they are making an iterative move in a longstanding game in which marketers have a long history of overstepping. In such cases, marketers would do well to employ a variation on TIT FOR TAT that has proven effective against the death spiral. It's called TIT FOR TWO TATS, and as the name implies, it allows the opponent to defect twice in a row before a retaliatory move, leaving more room for cooperation. Marketers

that feel stung by consumer backlash in the social media space should consider this strategy before backing away from the space.

3.10 The Marketer's Dilemma

At this point, if I have done my job, I have established that iterative games hold a great deal of hope for evolving toward cooperation, and that this may have some explanatory power for what's going in social media marketing. However, there's an important caveat: if you paid attention to the difference in the payoff table for marketers and the payoff table for the Prisoner's Dilemma, then you've noticed that the marketer's dilemma is a much tougher one than the prisoner's. Marketing has an equilibrium at the worst-case scenario (1-1, a death spiral), and a best-case scenario (4-4, a seemingly unattainable goal), whereas the Prisoner's Dilemma finds a point of equilibrium in the middle ground (neither the best or worst case).

So unlike the stable Prisoner's Dilemma, marketers and consumers vacillate back and forth over contested territory, each claiming conditional victories. Consumers gain temporary advantages with things like Tivo, which reduces their exposure to unwanted advertising; marketers regain the advantage with things like product placement in films, which replaces a portion of the lost exposure. And the dance continues. The game is perpetually suboptimal for both players, with the added stress of instability. So how could this game possibly evolve to the optimal equilibrium?

As the next chapter will show, the short answer is: Not every easily, and not all at once. The long answer may lie, at least in part, with the ethical dimension to Axelrod's analysis that goes beyond the material (or at least points-based) rewards that success in the Prisoner's Dilemma promises. As each side gains an understanding of the other's self-interest, something like empathy emerges, so that the act of cooperation is ennobling – thus changing the stakes of the game. Anyone who has developed brand loyalty based on a brand's apparent trustworthiness and care for its consumers has experienced this greater reward. Our loyalty in such cases goes beyond the purely rational; it touches an emotional core. Axelrod found that in these instances, "the very experience of sustained mutual cooperation altered the payoffs for both players, making mutual cooperation even more valued than it was before".

Flood and Dreshers' and Axelrod's experiments showed that sustained, stable cooperation is possible and even preferable in games for which the dominant strategy is to defect – certainly the marketing game is included in this category. Cooperation requires iteration, mutually agreed upon rules,

and most importantly, transparency that breeds trust, allowing all players to signal their willingness to cooperate. For marketers, this evolution will be a long journey, and failures tend to make headlines. But in the examples we'll examine in the next chapter, the worth and the long-term stability of cooperative strategies will prove themselves out.