

Strategic Moves

CHANGING THE GAME

Millions of people make at least one New Year's resolution every year. A Google search for the phrase "New Year's resolutions" produces 2.12 million pages. According to a U.S. government web site, the most popular of these resolutions is "lose weight." This is followed by "pay off debt," "save money," "get a better job," "get fit," "eat right," "get a better education," "drink less alcohol," and "quit smoking."¹

Wikipedia, the free online encyclopedia, defines a New Year's resolution as "a commitment that an individual makes to a project or a habit, often a lifestyle change that is generally interpreted as advantageous." Note the word "commitment." Most people have an intuitive understanding of it, in the sense of a resolve, a pledge, or an act of binding oneself. We will soon make the concept more precise in its gametheoretic usage.

What happens to all these wonderful life-improving plans? A CNN survey reports that 30 percent of the resolutions are not even kept into February, and only 1 in 5 stays on track for six months or longer.² Many reasons contribute to this failure: people set themselves excessively ambitious goals, they do not have good methods for measuring their progress, they lack the time, and so on. But by far the most important cause of failure is that, like Oscar Wilde, most people can resist anything except temptation. When they see and smell those steaks, french fries, and desserts, their diets are doomed. When those new electronic gadgets beckon, the resolution to keep the credit card in the wallet falters. When they are

sitting comfortably in their armchairs watching sports on TV, actually exercising seems too much like hard work.

Many medical and lifestyle advisers offer tips for success in keeping resolutions. These include basics such as setting reasonable and measurable goals, working toward them in small steps, setting up a regime of healthy food and exercise that is varied to prevent boredom, not getting discouraged, and not giving up after any setbacks. But the advice also includes strategies for creating the right incentives, and an important feature of these is a support system. People are advised to join groups that diet and exercise together and to publicize their resolutions among their family and friends. The feeling that one is not alone in this endeavor surely helps, but so does the shameful prospect of public failure.

This shame factor was powerfully harnessed by one of us (Nalebuff) in an ABC *Primetime* program, *Life: The Game*.³ As described in our opening chapter, the overweight participants agreed to be photographed wearing just a bikini. Anyone who failed to lose 15 pounds in the following two months would have his or her photographs displayed on national television and on the program's web site. The desire to avoid this fate served as a powerful incentive. All but one of the participants lost the 15 pounds or more; one failed but only narrowly.

Where does game theory come in? The struggle to lose weight (or to save more money) is a game of one's current self (who takes a long-run viewpoint and wants to improve health or wealth) against a future short-run self (who is tempted to overeat and overspend). The current self's resolution constitutes a commitment to behave better. But this commitment must be irreversible; the future self should be denied the possibility of reneging. The current self does this by taking an associated action—being photographed in an embarrassing outfit and giving up control over the use of these pictures to the program's producer to display if the weight loss is insufficient. This changes the game by changing the future self's incentives. The temptation to overeat or overspend still exists, but it is countered by the prospect of shameful exposure.

Such actions that change the game to ensure a better outcome for the player taking the actions are called *strategic moves*. In this chapter we will explicate and illustrate many of these moves. There are two aspects to consider: what needs to be done and how to do it. The former is amenable to the science of game theory, while the latter is specific to each situation—thinking up effective strategic moves in each specific context is more art than science. We will equip you with the basics of the science and try to convey some of the art through our examples. But we must leave it to you to further develop the art you will need in the games you play, based on your knowledge of the situations.

For our second example of changing the game, imagine yourself as an

American male teenager in the 1950s. You live in a small town. It is a clear Saturday evening. You are with a group of friends, playing games of rivalry to decide who is the alpha male. Tonight's contests start with a game of chicken. As you race toward a head-on collision, you know that the one who swerves first is the loser, or chicken. You want to win.

This is a dangerous game. If both of you attempt to win, both may end up in the hospital, or worse. We analyzed this game in chapter 4 from the perspective of Nash equilibrium (and in the context of the Stone Age hunters Fred and Barney) and found that it had two Nash equilibria, one where you go straight and your rival swerves, and the other where you swerve while your rival continues straight. Of course you prefer the first to the second. Here we take the analysis to a higher level. Can you do anything to achieve your preferred outcome?

You could establish a reputation as someone who never swerves. However, to do that you must have won similar games in the past, so the question just transfers itself to what you could have done in those games.

Here is a fanciful but effective device. Suppose you disconnect your steering wheel from the shaft and throw it out of the window in a manner that makes it very visible to your rival. He now knows that you *can't* swerve. The whole onus of avoiding the collision is on him. You have changed the game. In the new game you have just one strategy, namely to go straight. And then your rival's best (actually, least bad) response is to swerve. You are helpless as a driver, but that very helplessness makes you a winner in the game of chicken.

The way you have changed this game in your favor is surprising at first sight. By losing your steering wheel, you have restricted your own freedom of action. How can it be beneficial to have fewer choices? Because in this game, freedom to swerve is merely freedom to become the chicken; freedom to choose is freedom to lose. Our study of strategic moves will produce other seemingly surprising lessons.

This example also serves to motivate a fair warning about strategic moves. Their success is not guaranteed, and sometimes they can be outright dangerous. In reality, there are delays in action and observation. In chicken, what if your rival gets the same idea, and each of you simultaneously sees the other's steering wheel flying through the air? Too late. Now you are headed helplessly toward a crash.

So try these devices at your own risk, and don't sue us if you fail.

A LITTLE HISTORY

People and nations have made commitments, threats, and promises for millennia. They have intuitively recognized the importance of credibility in such actions. They have used such strategies and devised counterstrategies against other players' use of them. When Homer's Odysseus tied himself to the mast, he was making a credible commitment not to be lured by the song of the Sirens. Parents understand that while a threat of cold-bloodedly punishing a child for misbehavior is not credible, the threat "Do you want Mommy to get angry?" is much more believable. Kings throughout history have understood that voluntarily exchanging hostages—giving up a beloved child or other relative to live in a rival monarch's family—helps make their mutual promises of peaceful coexistence credible.

Game theory helps us understand and unify the conceptual framework of such strategies. However, in its first decade, game theory focused on characterizing different kinds of equilibria in a *given game*—backward reasoning in sequential-move games, minimax in two-person zero-sum games, and Nash in more general simultaneous-move games—and illustrating them in important contexts like the prisoners' dilemma, assurance, battle of the sexes, and chicken.⁴ Thomas Schelling gets the honor and credit for being the first person to develop the idea that one or both players might take actions to *change the game* as a central theme of game theory. His articles in the late 1950s and early 1960s, collected and elaborated on in his books *The Strategy of Conflict* (1960) and *Arms and Influence* (1965),⁵ gave us precise formulations of the concepts of commitment, threat, and promise. Schelling clarified just what is needed for credibility. He also analyzed the subtle and risky strategy of brinkmanship, which had previously been much misunderstood.

A more rigorous formal development of the concept of credibility, namely subgame perfect equilibrium, which is a generalization of the backward reasoning equilibrium we discussed in chapter 2, came a few years later from Reinhard Selten, who in 1994 was in the first group of game theorists to receive a Nobel Prize, jointly with John Nash and John Harsanyi.

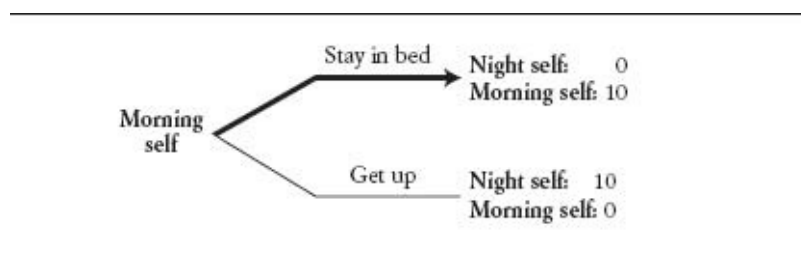
COMMITMENTS

Of course, you don't have to wait until New Year's Day to make a good resolution. Every night you can resolve to wake up early the next morning to get a good start on the day, or perhaps to go for that five-mile run. But you know that when morning comes, you will prefer to stay in bed for another half-hour or hour (or longer). This is a game of your resolute nighttime self against your own

future weak-willed morning self. In the game as structured, the morning self has the advantage of the second move. However, the nighttime self can change the game to create and seize first-mover advantage by setting the alarm clock. This is intended as a commitment to get out of bed when the alarm rings, but will it work? Alarm clocks have snooze buttons, and the morning self can hit the button, repeatedly. (Of course an even earlier self could have searched and bought an alarm clock without a snooze button, but that may not have been possible.) The night self can still make the commitment credible by keeping the alarm clock on the wardrobe across the room instead of on the bedside table; then the morning self will have to get out of bed to shut off the noise. If this is not enough and the morning self stumbles straight back into bed, then the night self will have to think up some other device, perhaps an alarm clock that at the same time starts brewing coffee, so the wonderful smell will induce the morning self out of bed.*

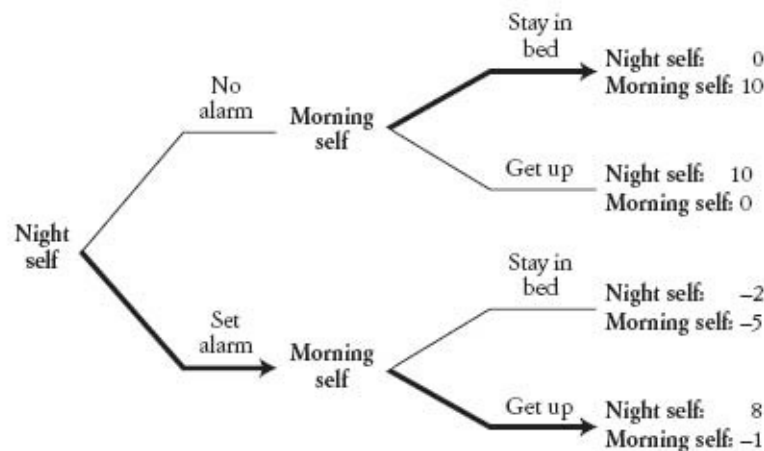
This example nicely illustrates the two aspects of commitments and credibility: what and how. The “what” part is the scientific or gametheoretic aspect—seizing first-mover advantage. The “how” part is the practical aspect or the art—thinking up devices for making strategic moves credible in a specific situation.

We can illustrate the mechanics or science of the commitment of the alarm clock using the tree diagrams of chapter 2. In the original game, where the night self takes no action, the game is trivial:



The morning self stays in bed and gets its preferred payoff, which we have assigned 10 points, leaving the night self with its worse payoff, to which we have assigned 0 points. The precise number of points does not matter much; all that matters is that for each self, the preferred alternative is assigned more points than the less-preferred one.

The night self can change the game into the following:



Now the payoff numbers matter a little bit and need more explanation. Along the upper main branch where the night self does not set an alarm, the tree is as before. Along the lower main branch, we have supposed that the night self has a small cost, which we have set at 2 points, of setting the alarm clock. So if the morning self heeds the alarm and gets up, the night self will get 8 points, instead of the 10 in the original game. But if the morning self were to ignore the alarm, the night self would get -2 points since the cost of setting the alarm was wasted. The morning self has an annoyance cost of hearing the alarm; it is only 1 if it gets out of bed to turn off the alarm quickly but would be intolerably large (15 points) if it stayed in bed and the alarm went on and on, converting the pleasure of the bed (10) into a payoff of -5 (= 10-15). If the alarm has been set, the morning self prefers -1 to -5 and gets up. The night self looks ahead to this, and reasons that setting the alarm will give it 8 points in the eventual outcome, which is better than the zero it would get in the original game.* Therefore, in the backward reasoning equilibrium of the game, the morning self does get up if an alarm has been set, and the night self sets the alarm.

A more striking aspect of commitment may be seen if we represent this game in a game table, instead of a tree.

		Morning self	
		Stay in bed	Get up
Night self	No alarm	0 10	10 0
	Set alarm	-2 -5	8 -1

The table shows that for each given strategy of the morning self, the night self's payoff from Set alarm is smaller than that from No alarm: -2 is less than 0 , and 8 is less than 10 . Therefore for the night self, the strategy Set alarm is dominated by No alarm. Nevertheless, the night self finds it desirable to commit to Set alarm!

How can it be good to choose a dominated strategy and not play a dominant strategy? To understand this, we need to understand the concept of dominance more clearly. No alarm dominates Set alarm from the perspective of the night self because, for each *given* strategy of the morning self, No alarm yields a higher payoff to the night self than does Set alarm. If the morning self chooses Stay in bed, the night self gets 0 from No alarm and -2 from Set alarm; if the morning self chooses Get up, the night self gets 10 from No alarm and 8 from Set alarm. If moves are simultaneous, or if the night self moves second, he cannot affect what the morning self chooses and must accept it as given. But the very purpose of a strategic move is to *alter* the other player's choice, not to take it as given. If the night self chooses Set alarm, the morning self will choose Get up and the night self will have payoff 8 ; if the night self chooses No alarm, the morning self will choose Stay in bed and the night self's payoff will be 0 ; and 8 is greater than 0 . The payoffs of 10 and -2 , and their comparisons with 8 and 0 , respectively, become irrelevant. Thus the concept of dominance loses its significance for a first mover in a sequential game.

For most of the examples we give in this chapter, you will be able to get the idea without drawing any such explicit trees or tables, so we will generally offer only verbal statements and reasoning. But you can reinforce your understanding of the game, and of the tree method, by drawing them for yourself if you wish.

THREATS AND PROMISES

A commitment is an *unconditional* strategic move; as the Nike slogan says, you "just do it"; then the other players are followers. The night self simply sets the alarm on the bureau and the timer on the coffee machine. The night self has no further moves in the game; one might even say that the night self ceases to exist in the morning. The morning self is the follower player or second mover, and its best (or least bad) response to the night self's commitment strategy is to get out of bed.

Threats and promises, on the other hand, are more complex *conditional* moves; they require you to fix in advance a *response rule*, stating how you

would respond to the other player's move in the actual game. A threat is a response rule that punishes others who fail to act as you would like them to. A promise is an offer to reward other players who act as you would like them to.

The response rule prescribes your action as a response to the others' moves. Although you act as a follower in the actual game, the response rule must put be in place *before* others make their moves. A parent telling a child "No dessert unless you eat your spinach" is establishing such a response rule. Of course, this rule must be in place and clearly communicated before the child feeds her spinach to the dog.

Therefore such moves require you to change the game in more complex ways. You must seize the first-mover status in the matter of putting the response rule in place and communicating it to the other player. You must ensure that your response rule is credible, namely that if and when the time comes for you to make the stated response, you will actually choose it. This may require changing the game in some way to ensure that the choice is in fact best for you in that situation. But in the game that follows, you must then have the second move so you will have the ability to respond to the other's choice. This may require you to restructure the order of moves in the game, and that adds its own difficulties to your making the strategic move.

To illustrate these ideas, we will use the example of the pricing rivalry of the catalog merchants B. B. Lean and Rainbow's End, which we developed as a simultaneous-move game in chapters 3 and 4. Let us recapitulate its basic points. The two are competing over a specific item, a deluxe chambray shirt. Their joint interests are best served if the two collude and charge a monopoly price of \$80. In this situation each will make a profit of \$72,000. But each has the temptation to undercut the other, and if they both do so, in the Nash equilibrium each will charge only \$40 and make a profit of only \$40,000. This is their prisoners' dilemma, or a lose-lose game; when each gives way to the temptation to make a bigger profit for itself, both lose.

Now let us see if strategic moves can resolve the dilemma. A commitment by one of them to keep its price high won't do; the other will simply exploit it to the disadvantage of the first. What about conditional moves? Rainbow's End might employ a threat ("If you charge a low price, so will I") or a promise ("If you keep your price at the monopoly level, so will I"). But if the actual game of choosing prices in the catalog has simultaneous moves in the sense that neither can observe the other's catalog before setting its own in print, how can Rainbow's End *respond* to B. B. Lean's move at all? It must change the game so it has the opportunity to choose its price after it knows the other's price.

A clever commonly used device, the meet-the-competition clause, achieves

this purpose. In its catalog, Rainbow's End prints the price \$80, but with a footnote: "We will meet any lower price charged by any competitor." Now the catalogs are printed and mailed simultaneously, but if B. B. Lean has cheated and printed a price lower than \$80, perhaps all the way down to the Nash equilibrium price of \$40, then Rainbow's End automatically matches that cut. Any customer who might have a slight preference or loyalty toward Rainbow's End need not switch to B. B. Lean for its lower price, he can simply order from Rainbow's End as usual and pay the lower price listed in the B. B. Lean catalog.

We will return to this example again to illustrate other aspects of strategic moves. For now, just note two distinct aspects: the scientific or "what" aspect (the threat to match any price cut) and the art or the "how" aspect (the meet-the-competition clause that makes the threat possible and credible).

DETERRENCE AND COMPELLENCE

The overall purpose of threats and promises is similar to that of commitments, namely, to induce the others to take actions different than they would otherwise. In the case of threats and promises, it is useful to classify the overall purpose into two distinct categories. When you want to stop the others from doing something they would otherwise do, that is *deterrence*. Its mirror image, namely to compel the others to do something they would not otherwise do, can then be termed *compellence*.⁶

(MINI) TRIP TO THE GYM NO. 4

Set up the tree for the game of the cold war, and show how the U.S. threat changes the equilibrium outcome of the game.

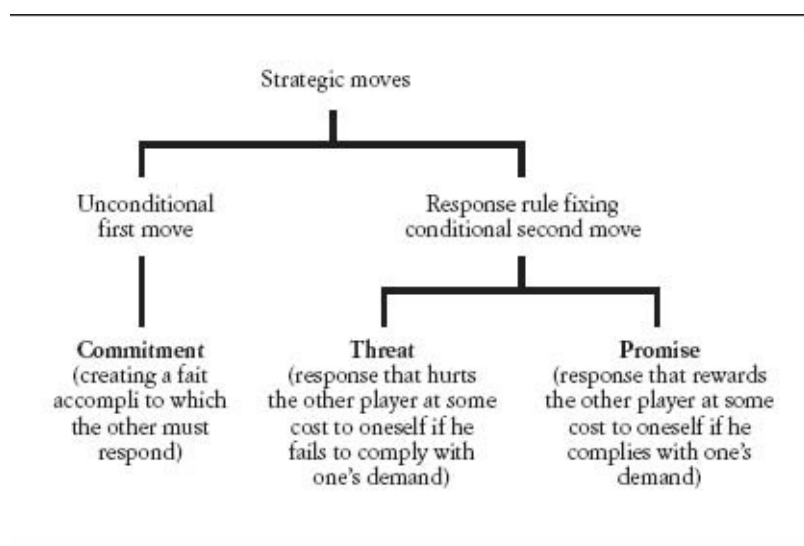
When a bank robber holds the employees hostage and establishes a response rule that he will kill them if his demands are rejected, he is making a compellent threat. When, during the cold war, the United States threatened to respond with nuclear weapons if the Soviet Union attacked any NATO country, it made a deterrent threat. The two threats share a common feature: *both* sides will bear an extra cost if the threat has to be carried out. The bank robber compounds the punishment he will face when caught if he adds murder to his original crime of armed robbery; the United States would suffer horribly in a nuclear war when it

could have lived with a Soviet-dominated Europe.

Promises can also be compellent or deterrent. A compellent promise is designed to induce someone to take a favorable action. For example, a prosecutor who needs a witness to buttress his case promises one defendant a more lenient sentence if he turns state's evidence against his codefendants. A deterrent promise is designed to prevent someone from taking an action that is against your interests, as when the mobsters promise a confederate they will protect him if he keeps his mouth shut. Like the two kinds of threats, the two promises also share a common feature. After the other player has complied with one's wishes, the promisor no longer needs to pay the cost of delivering the reward and has the temptation to renege. Thus, after the mob bosses on trial are acquitted for lack of evidence, they might kill the confederate anyway to avoid the risk of any future trouble or blackmail.

A QUICK REFERENCE GUIDE

We have thrown many concepts at you thick and fast. To help you remember them, and to refer to them later at a glance, here is a chart:



And here is a table summarizing, in the form of pregame statements of the strategic mover, how threats and promises seek to achieve each of the two aims: deterrence and compellence. “If, in the game to follow, you...

	Deterrence	Compellence
Threat	...do what I don't want you to do...	...don't do what I want you to do...
	...then I will respond with an action which will hurt you (and will also hurt me).	
Promise	...don't do what I don't want you to do...	...do what I want you to do...
	...then I will respond with an action which will reward you (and will be costly for me)."	

WARNINGS AND ASSURANCES

All threats and promises have a common feature: the response rule requires you to take actions that you would not take in its absence. If instead the rule merely says that you will do what is best at the time, this is as if there is no rule: there is no *change* in others' expectations about your future actions and hence no change in their actions. Still, there is an informational role for stating what will happen, even without any rule; these statements are called *warnings* and *assurances*.

When it is in your interest to carry out a "threat," we call this a *warning*. For example, if the president warns he will veto a bill not to his liking, this is simply an indication of his intentions. It would be a threat if he were willing to sign the bill but strategically committed to veto it in order to induce Congress to offer something even better.

To illustrate this in a business context, let us examine whether B. B. Lean's matching Rainbow's End's price cuts constitutes a threat or a warning. In chapter 4 we considered the best response of B. B. Lean to various prices that Rainbow's End could conceivably charge. We found that it was somewhere between zero and full response. If B. B. Lean were to keep its price unchanged while Rainbow's End cut its price, then B. B. Lean would lose too many customers to its rival. But if B. B. Lean were to match Rainbow's End's price cut dollar for dollar, its own profit margin would be squeezed too far. In the example we developed, B. B. Lean struck the optimal balance between these two considerations by reducing its price by 40 cents for each dollar reduction in Rainbow's End's price.

But if B. B. Lean wants to threaten Rainbow's End to deter it from initiating any price cuts, it may need to threaten a larger response than the 40 cents per dollar that would be optimal in the event of an actual price cut by Rainbow's

End. In fact, B. B. Lean may want to threaten a superaggressive response of more than a dollar. It can do so by printing a beat-the-competition clause instead of merely a meet-the-competition clause in its catalog. Such devices are genuine threats in our terminology. B. B. Lean would find it costly to carry out the actions if put to the test by Rainbow's End. Its threat is made credible by printing its policy in the catalog, so its customers can rely upon it as law, and B. B. Lean cannot renege on it. If B. B. Lean had said in its catalog: "For every dollar that Rainbow's End's price falls short of \$80, we will charge 40 cents less than our catalog price of \$80," this would be merely a warning to Rainbow's End; if put to the test, B. B. Lean would want to go through with the stated response anyway.

When it is in your interest to carry out a promise, we call this an assurance. In the shirt pricing example, B. B. Lean may secretly want to tell Rainbow's End that if they hold to the collusive price of \$80, so will B. B. Lean. In the game played once, this is not in the interest of B. B. Lean after the fact. Therefore it is a genuine strategic move, namely a promise. If the game was repeated so that continued mutual cooperation was an equilibrium, as we saw in chapter 3, then the statement from B. B. Lean would be an assurance, intended merely to inform Rainbow's End that B. B. Lean was quite aware of the nature of the repeated game and how it offered a resolution to their prisoners' dilemma.

To reiterate the point, threats and promises are truly strategic moves, whereas warnings and assurances play more of an informational role. Warnings or assurances do not change your response rule in order to influence another party. Instead, you are simply informing them of how you will want to respond based on their actions. In stark contrast, the sole purpose of a threat or promise is to change your response rule away from what will be best when the time comes, not in order to inform but to manipulate.

Because threats and promises indicate that you will act against your own interest, their credibility becomes the key issue. After others have moved, you have an incentive to break your threat or promise. Some other accompanying change in the game is needed to ensure credibility. Without credibility, other players will not be influenced by mere words. Children who know that their parents get pleasure from giving them toys are not influenced by threats to withhold toys unless the parents take some prior action to make the threat credible.

Strategic moves, therefore, contain two elements: the planned course of action and the associated actions that make this course credible. We will try to give you a better appreciation of both aspects by making two passes through the ideas. In the remainder of this chapter we focus attention on the former, or what

needs to be done to make threats and promises. Think of this as a menu of moves. In the next chapter we will turn our focus to the recipes for credibility—that is, how to make threats and promises believable and therefore effective.

OTHER PLAYERS' STRATEGIC MOVES

It is natural to think of the advantages you can get from strategic moves, but you should also think about how such moves made by other players will affect you. In some cases you might even benefit by relinquishing the opportunity to make such a move and purposely allowing someone else to do so. Three such logical possibilities are:

You may allow someone to make an unconditional move before you respond.

You may wait for a threat before taking any action.

You may wait for a promise before taking any action.

We have already seen examples in which someone who could move first does even better by relinquishing this option, allowing the other side to make an unconditional move. This is true whenever it is better to follow than to lead, as in the chapter 1 tale of the America's Cup race (and in the chapter 14 case study on gambling at the Cambridge May Ball). More generally, if the game, played sequentially, has a second mover advantage, you can benefit by arranging things so that the other player must move first, thereby making an unconditional commitment. While it can be advantageous to give up the initiative, this is not a general rule. Sometimes your goal will be to prevent your opponent from making an unconditional commitment. This was the motivation behind Chinese military strategist Sun Tzu's advice to leave the enemy an escape route—the idea is to prevent the enemy from making a commitment to fight to the death.

It is never advantageous to allow others to threaten you. You could always

do what they wanted you to do without the threat. The fact that they can make you worse off if you do not cooperate cannot help, because it limits your available options. But this maxim applies only to allowing threats. If the other side can make promises, then you can both be better off. A simple example is the prisoners' dilemma, where both players can benefit if even one player has some way to make a credible promise to keep quiet. Note that it must be a conditional move, a promise, not an unconditional commitment. If the other player were to make a commitment to keep quiet, you would simply exploit it by confessing, and, knowing this, he would not make such a move.

SIMILARITIES AND DIFFERENCES BETWEEN THREATS AND PROMISES

Sometimes the distinctions between threats and promises are blurred. A friend was mugged in New York City with the following promise: If you "lend" me twenty dollars, I promise I won't hurt you. More relevant was the mugger's implicit threat that if our friend *didn't* lend him the money, he would be hurt.

As this story suggests, the distinction between a threat and a promise depends only on what you call the status quo. The traditional mugger threatens to hurt you if you don't give him some money. If you don't, he starts hurting you, making that the new status quo, and promises to stop once you give him money. A compellent threat is just like a deterrent promise with a change of status quo; likewise, a deterrent threat and a compellent promise differ only in their status quo.

So should you use a threat or a promise? The answer depends on two considerations. The first is the cost. A threat can be less costly; in fact, it is costless if it is successful. If it changes the other player's behavior in the way you want, you don't have to carry out the costly action you had threatened. A promise, if successful, must be fulfilled—if the other player acts as you want him to, you have to deliver the costly action you had promised. If a company could threaten its employees with terrible consequences should their performance fall short of being excellent, it could save a lot of money that it usually pays out to fulfill its promises of incentive bonuses. Indeed, Stalin tried using just sticks instead of carrots—threats of being sent to the Gulag in Siberia instead of promises of better pay or living conditions—to get good performance from Soviet workers. But his system did not work because its methods for judging performance were inaccurate, arbitrary, and corrupt. We will return to this point in the next section.

The second consideration in the choice between a threat and a promise is whether the purpose is deterrence or compellence. The two have different time dimensions. Deterrence does not necessarily have a deadline. It simply involves telling the other player not to do such and such, and credibly communicating the bad consequences that would follow if he takes the forbidden action. So the United States says to the Soviet Union: “Don’t invade Western Europe,” or God says to Adam and Eve, “Don’t eat the apple.” “When?” “Ever.”* Therefore deterrence can be achieved more simply and better by a threat. You set up a tripwire, and it is up to the other to decide whether to trigger it.

In contrast, compellence must have a deadline. When a mother says to her child, “Clean your room,” a time limit such as “before 5:00 P.M. today” must accompany it. Otherwise the child can defeat the purpose by procrastination: “I have soccer practice today; I will do it tomorrow,” and when tomorrow comes, some other, more urgent task will come up. If the mother has threatened some dire consequence, she does not want to invoke it for each seemingly tiny delay. The child can defeat her threat “slice by slice,” a strategy that Schelling calls *salami tactics*.

Therefore compellence is often better achieved by giving the other player the incentive not to procrastinate. This means that earlier performance must get a better reward or lighter punishment. This is a promise. The mother says: “You will get that special treat for dessert when you have cleaned your room,” and the mugger says: “The knife at your throat will go away as soon as you have given me your money.”

CLARITY AND CERTAINTY

When making a threat or a promise, you must communicate to the other player quite clearly what actions will bring what punishment (or what reward). Otherwise, the other may form a wrong idea of what is forbidden and what is encouraged and miscalculate the consequences of his actions. Stalin’s stick-type “incentives” for workers in the Soviet Union suffered from this crucial flaw. The monitoring system was arbitrary and corrupt, so the worker stood almost as much a risk of going to Siberia if he worked hard as if he shirked. So why work?

But clarity does not have to be a simple either-or choice. In fact, such a stark alternative may be poor strategy. The United States wanted to deter the Soviet Union from invading Western Europe. But threatening all-out nuclear war in the event of the smallest transgression, say a handful of soldiers straying across the border, might be too risky. When a company wants to promise rewards to its

workers for improved productivity, a bonus that increases gradually with an increase in output or profit may be better than offering nothing if the performance does not exceed a set target, and a very large sum if it does.

For a threat or promise to have its desired effect, the other player must believe it. Clarity without certainty doesn't cut it. Certainty does not mean a complete lack of risk. When a company offers stock bonuses to its managers, the value of the promised reward is uncertain, influenced by many factors that affect the market and are outside the control of the manager. But the manager should be told just how many shares he will get in relation to the immediately measurable indicator of his performance on which the bonus is based.

Nor does certainty require that everything happens at once. Threats and promises that work in multiple small steps are especially useful against salami tactics. When we give exams to students, there are always a few who attempt to keep writing after the time is up, in the hope of getting a few extra points. Grant them an extra minute and they will go past that, grant another minute and it becomes five, and so on. The dire punishment of refusing to accept an exam that is two or three minutes late would not be credible, but levying a graduated penalty of a few grade points per minute of delay is perfectly credible.

LARGE THREATS

If a threat is successful, the threatened action does not have to be carried out. Even though it may be costly for you to carry it out, since you don't have to do so, the cost is irrelevant. So why not use a huge threat that would really frighten the other player into acceding to your wishes? Instead of politely asking your dinner table neighbor to please pass the salt, why don't you threaten him with: "If you don't pass the salt, I will smash your head"? Instead of patiently negotiating with trading partner countries in an attempt to persuade them to lower their barrier against our exports, why doesn't the United States threaten that if they don't buy more of our beef or wheat or oranges, we will nuke them?

This is an obviously horrific idea; the threats are too large to be useable or believable. In part this is because they would generate terror and revulsion at the gross violation of all social norms of behavior. But in part it is also because the assumption that you would never have to carry out the threatened action is not 100 percent valid. Suppose something goes wrong. Your dinner table neighbor may be the obstinate kind who revolts at any prospect of bullying, or a tough guy who enjoys an opportunity for a fight. If he refuses to comply, you must either go through with the threatened action or back down and face the humiliation and

loss of reputation. Similar considerations apply to the United States if it tries to threaten another country with a harsh military action in an economic dispute. Even slight risks of such hugely costly errors provide strong arguments for keeping threats at the smallest level needed to keep them effective.

Very often you don't know the exact size of a threat that is needed to deter or compel your adversary. You want to keep the size as low as possible to minimize the cost to you in the event that things go wrong and you have to go through with the action. So you start small and gradually raise the size of the threat. This is the delicate strategy of brinkmanship.

BRINKMANSHIP

In the book and movie *L.A. Confidential*, the “good cop” Ed Exley is interrogating a suspect, Leroy Fontaine, when the hot-tempered cop Bud White intervenes:

The door banged open. Bud White stepped in, threw Fontaine against the wall.

Ed froze.

White pulled out his .38, broke the cylinder, dropped shells on the floor. Fontaine shook head to toe; Ed kept freezing. White snapped the cylinder shut, stuck the gun in Fontaine's mouth. “One in six. Where's the girl?”

Fontaine chewed steel; White squeezed the trigger twice: clicks, empty chambers. [So now the risk has risen to one in four.] Fontaine slid down the wall; White pulled the gun back, held him up by his hair. “*Where's the girl?*”

Ed kept freezing. White pulled the trigger—another little click. [So now it is one in three.] Fontaine, bug-eyed. “S-ss-sylvester F-fitch, one-o-nine and Avalon, gray corner house please don’ hurt me no—”

White ran out.⁷

Obviously White is threatening Fontaine to compel him to reveal the information. But what is the threat? It is not simply: “If you don’t tell me, I will kill you.” It is: “If you don’t tell me, I will pull the trigger. If the bullet happens to be in the firing chamber, you will die.” It is creating a *risk* that Fontaine will be killed. And every time the threat is repeated, the risk is increasing. Finally, at one in three, Fontaine finds the risk too high and spills the information. But there were other possibilities: White may have feared that the information would die with Fontaine, found the risk too high, backed down, and tried something else. Or the thing they both feared—the bullet reaches the firing chamber and Fontaine dies—might have come about.

A similar situation arises in the movie *The Gods Must Be Crazy*. There has been an unsuccessful attempt on the life of the president of a country in Africa. The presidential guards have caught one of the attackers, and he is being interrogated for information about the rest of his group. He stands blindfolded with his back to the open door of a helicopter with its rotors whirring. The officer facing him asks: “Who is your leader? Where is your hideout?” No answer. The officer pushes him out of the door of the helicopter. The scene switches to the outside. We see that the helicopter is actually hovering just a foot off the ground, and the man has fallen on his back. The interrogating officer appears at the door, laughs, and says to the man: “The next time it will be a little bit higher.” The scared man gives away the information.

What is the purpose of such threats of increasing risk? We argued in the previous section that there are good reasons for keeping the size of a threat down to the smallest level that will have the desired effect. But you may not know in advance the smallest effective size of a threat. That is why it makes sense to start small, and increase it gradually to find out when it works. As the size of the threatened action grows, the cost of carrying it out grows too. In the examples above, the way to increase the size of the threat is to increase the *risk* of the bad thing happening. The maker and the recipient of the threat are then engaged in a

game of exploring each other's tolerance for this cost or the risk. Is a one-in-four chance of Fontaine being killed too large for Fontaine or for White? If not, then try one in three. They continue this eyeball-to-eyeball confrontation until one of them blinks—or until the outcome they both fear transpires.

This is the strategy that Schelling called *brinkmanship*.^{*} The term is often interpreted as taking an adversary to the brink of disaster in order to get him to blink first. Standing on the brink, you threaten to push him off if he fails to comply with your wishes. Of course, he will take you down with him. That is why, says Schelling, the pure and simple threat of cold-bloodedly pushing the adversary off the brink is not credible.

If the brink is clearly marked and provides a firm footing, no loose pebbles underfoot and no gusts of wind to catch one off guard, if each climber is in full control of himself and never gets dizzy, neither can pose any risk to the other by approaching the brink.... [W]hile either can deliberately jump off, he cannot credibly pretend that he is about to. Any attempt to intimidate or to deter the other climber depends on the threat of slipping or stumbling.... [O]ne can credibly threaten to fall off accidentally by standing near the brink.

Deterrence has to be understood in relation to this uncertainty....A response that carries some risk of war [through a compounding of actions and reactions, of calculations and miscalculations, of alarms and false alarms] can be plausible, even reasonable, at a time when a final, ultimate decision to have a general war would be implausible or unreasonable.⁸

The 1962 Cuban missile crisis provided perhaps the most famous example of brinkmanship. The Soviet Union, under its mercurial leader Nikita Khrushchev, had begun to install nuclear missiles on Cuba, ninety miles from the American mainland. On October 14, American reconnaissance airplanes brought back photographs of missile sites under construction. After a week of tense discussions within his administration, on October 22 President John F. Kennedy announced a naval quarantine of Cuba. Had the Soviet Union taken up the challenge, the crisis could have escalated to the point of all-out nuclear war between the superpowers. Kennedy himself estimated the probability of this as

“between one out of three and even.” But after a few anxious days of public posturing and secret negotiation, Khrushchev looked over the nuclear brink, did not like what he saw, and pulled back. In return for a face-saving compromise involving eventual withdrawal of U.S. missiles from Turkey, he ordered the Soviet missiles in Cuba dismantled and shipped back.⁹

Just where was the brink in the Cuban missile crisis? Had the Soviets tried to defy the blockade, for example, the United States was unlikely to launch its strategic missiles at once. But events and tempers would have heated up another notch, and the risk of Armageddon would have increased perceptibly.

Soldiers and military experts speak of the “fog of war”—a situation in which both sides act with disrupted lines of communication, individual acts of fear or courage, and a great deal of general uncertainty. There is too much going on to keep everything under control. This serves the purpose of creating some risk. Even the president found it difficult to control the operations of the naval blockade of Cuba once put into play. Kennedy tried to move the blockade from 800 miles to 500 miles off the shore of Cuba in order to give Khrushchev more time. Yet evidence based on the first ship boarded, the *Marcula* (a Lebanese freighter under charter by the Soviets), indicates that the blockade was never moved.¹⁰

The key to understanding brinkmanship is to realize that the brink is not a sharp precipice but a slippery slope, getting gradually steeper. Kennedy took the world some way down this slope; Khrushchev did not risk going farther, and then the two arranged a pullback to the safe ground above.*

The essence of brinkmanship is the deliberate creation of risk. This risk should be sufficiently intolerable to your opponent to induce him to eliminate the risk by following your wishes. The game of chicken, discussed in the preceding chapters, is of this kind. Our earlier discussions supposed that each driver had just two choices, whether to swerve or to go straight. But in reality the choice is not whether to swerve but when to swerve. The longer the two keep on going straight, the greater the risk of a collision. Eventually the cars get so close to each other that even if one of the drivers decides that the danger is too high and swerves, it may be too late to avoid a collision. In other words, brinkmanship is “chicken in real time”: a game of increasing risk, just like the interrogation games in the movies.

Once we recognize this, we see brinkmanship everywhere. In most confrontations, for example, between a company and a labor union, a husband and a wife, a parent and a child, and the president and Congress, one or both of the players cannot be sure of the other party’s objectives and capabilities.

Therefore most threats carry a risk of error, and almost every threat contains an element of brinkmanship. Understanding the potentialities and risks of this strategic move can prove crucial in your life. Use it carefully, and understand that even with the best care it may fail, because the bad thing you and the other player both dread may come to pass while you are raising the stakes. If your assessment is that in this confrontation you will “blink first”—that is, the probability of the bad thing happening will get too large for your own tolerance before the limit of the other player’s tolerance is reached—then you may be better advised not to embark on the path of brinkmanship in the first place.

We will return to some aspects of the art of practicing brinkmanship in the next chapter. For now, we end on a cautionary note. With any exercise of brinkmanship, there is always the danger of falling off the brink. While we look back at the Cuban missile crisis as a successful use of brinkmanship, our evaluation would be very different if the risk of a superpower war had turned into a reality. The survivors would have cursed Kennedy for recklessly and unnecessarily flaming a crisis into a conflagration. Yet in any exercise of brinkmanship, the risk of falling off the brink can turn into a reality. The massacre of the Chinese students occupying Beijing’s Tiananmen Square in June 1989 is a tragic example. The students were on a collision course with the hard-liners in their government. One side would have to lose; either the hard-liners would cede power to more reform-minded leaders or the students would compromise on their demands. During the confrontation, there was a continual risk that the hard-liners would overreact and use force to squelch the democracy movement. When two sides are playing a game of brinkmanship and neither side is backing down, there is a chance that the situation will get out of control, with tragic consequences.

In the aftermath of Tiananmen Square, government leaders became more aware of the dangers in brinkmanship—for both sides. Faced with similar democracy protests in East Germany and Czechoslovakia, the communist leaders decided to give in to popular demands. In Romania, the government tried to hold firm against a reform movement, using violent repression to maintain power. The violence escalated almost to the level of a civil war, and in the end President Nicolae Ceaușescu was executed for crimes against his people.

CASE STUDY: TWO WRONGS KEEP THINGS RIGHT

Parents often face a difficult problem in punishing their children for bad behavior. Children have an uncanny sense of when the parents’ threat to punish

may not be credible. They recognize that the punishment may hurt the parents as much as the children (although for different reasons). The standard parental dodge to this inconsistency is that the punishment is for the child's own good. How can parents do a better job at making their threat to punish bad behavior credible?

Case Discussion

With two parents and one child, we have a three-person game. Teamwork can help the parents make an honest threat to punish a misbehaving child. Say the son misbehaves, and the father is scheduled to carry out the punishment. If the son attempts to rescue himself by pointing out the "irrationality" of his father's actions, the father can respond that he would, given the choice, prefer not to punish his son. But, were he to fail in carrying out the punishment, that would be breaking an agreement with his wife. Breaking that agreement would be worse than the cost of punishing the child. Thus the threat to punish is made credible.

Single parents can play this game, but the argument gets much more convoluted, as the punishment agreement must be made with the child. Now if the son attempts to rescue himself by pointing out the "irrationality" of his father's actions, the father can respond that he would, given the choice, prefer not to punish his son. But, were he to fail in carrying out the punishment, then this would be a misdeed on his part, a misdeed for which he should be punished. Thus, he is punishing his son only to prevent getting punished himself. But who is there to punish him? It's the son! The son replies that were his father to forgive him, he too would forgive his father and not punish his father for failing to punish him. The father responds that were his son to fail to punish him for being lenient, this would be the *second* punishable offense done by the son in the same day! And so on and so forth do they keep each other honest. This may seem a little far-fetched, but it is no less convoluted than most real arguments used to justify punishing kids who misbehave.

A compelling example of how two people can keep each other honest comes from Yale economist Dean Karlan. Dean was keen to lose weight and so wrote a contract with one of his friends that if either of them was ever above 175 pounds, the overweight one would owe the other \$1,000 per pound. Dean is a professor, and so that was a large financial penalty looming over his head. The threat worked for him and for his friend, too. But there was always the question of whether the friends would actually take each other's money.

Dean's friend got lazy and crept up to 190. Dean called him on the scale and took \$15,000 of his money. Dean didn't want to take money from his friend, but he knew that by doing so, his friend would then have no hesitation to take the money back should Dean ever fail. Dean engaged in the punishment to ensure that he would be punished if need be. Knowing that this threat is real has worked for Dean. If you'd like, he offers this service to others through his Commitment Store, which we discuss in the next chapter.

This concludes our brief sketch of the "what" of threats and promises. (For more practice, have a look at the "Arms Across the Ocean" case study in chapter 14.) Although we did have to say something about credibility, that was not the focus so far. In the next chapter we turn our attention to the issue of making credible strategic moves. We can offer only a general guide to this; it is largely an art that you must acquire by thinking about and through the dynamics of your own specific situation.