Deterrence and Compellence

THOMAS C. SCHELLING

THE DISTINCTION BETWEEN DETERRENCE AND “COMPELLENCE”

Blockade illustrates the typical difference between a threat intended to make an adversary do something and a threat intended to keep him from starting something. The distinction is in the timing and in the initiative, in who is to make the first move, in whose initiative is put to the test. To deter an enemy’s advance it may be enough to burn the escape bridges behind me, or to rig a trip-wire between us that automatically blows us both up when he advances. To compel an enemy’s retreat, though, some threat of engagement, I have to be committed to move. (This requires setting fire to the grass behind me as I face the enemy, with the wind blowing toward the enemy.) I can block your car by placing mine in the way; my deterrent threat passive, the decision to collide is up to you. But if you find me in your way and threaten to collide unless I move, you enjoy such advantage; the decision to collide is still yours, and I still enjoy deterrence. You have to arrange to have to collide unless I move, and that is a degree more complicated. You have to get up so much speed that you cannot stop in time and that only I can avert the collision; this may not be easy. If it takes more time to start a car than to stop one, you may be unable to give me the “last clear chance” to avoid collision by vacating the street.

The threat that compels rather than deters often requires that the punishment be administered until the other acts, rather than if he acts. This is because often the only way to become committed to an action is to initiate it. This means, though, that [action initiated has to be tolerable to the initiator, and tolerable over whatever period of time is required for the pressure work on the other side.] For deterrence, the trip-wire can threaten to blow things up out of all proportion to what is being protected, because if the threat works the thing never goes off. But to hold a large bomb and threaten to throw it unless somebody moves cannot work so well; the threat is not believable until the bomb is actually thrown and by then the damage is done.

There is, then, a difference between deterrence and what we might, for want of a better word, call compellence. The dictionary’s definition of “deter” corresponds to contemporary usage: to turn aside or discourage through fear; hence, to prevent from action by fear of con-
Deterrence and compellence

sequences. A difficulty with our being an unaggressive nation, one whose announced aim has usually been to contain rather than to roll back, is that we have not settled on any conventional terminology for the more active kind of threat. We have come to use “defense” as a euphemism for “military,” and have a Defense Department, a defense budget, a defense program, and a defense establishment; if we need the other word, though, the English language provides it easily. It is “offense.” We have no such obvious counterpart to “deterrence.” “Coercion” covers the meaning but unfortunately includes “deterrent” as well as “compellent” intentions. “Intimidation” is insufficiently focused on the particular behavior desired. “Compulsion” is all right but its adjective is “compulsive,” and that has come to carry quite a different meaning. “Compellence” is the best I can do.

Deterrence and compellence differ in a number of respects, most of them corresponding to something like the difference between statics and dynamics. Deterrence involves setting the stage—by announcement, by rigging the trip-wire, by incurring the obligation—and waiting. The overt act is up to the opponent. The stage-setting can often be nonintrusive, nonhostile, nonprovocative. The act that is intrusive, hostile, or provocative is usually the one to be deterred; the deterrent threat only changes the consequences if the act in question—the one to be deterred—is then taken. [Compellence, in contrast, usually involves initiating an action (or an irrevocable commitment to action) that can cease, or become harmless, only if the opponent responds.] The overt act, the first step, is up to the side that makes the compellent threat. To deter, one digs in, or lays a minefield, and waits—in the interest of inaction. To compel, one gets up enough momentum (figuratively, but sometimes literally) to make the other act avoid collision.

Deterrence tends to be indefinite in its timing. “If you cross the line we shoot in self-defense, or the mines explode.” When? Whenever you cross the line—preferably never, but the timing is up to you. If you cross it, then is when the threat is fulfilled, either automatically, if we’ve rigged it so, or by obligation that immediately becomes due. But we can wait—preferably forever; that’s our purpose.

Compellence has to be definite: We move, and you must get out of the way. By when? There has to be a deadline, otherwise tomorrow never comes. If the action carries no deadline it is only a posture, or a ceremony with no consequences. If the compellent advance is like Zeno’s tortoise that takes infinitely long to reach the border by traversing, with infinite patience, the infinitely small remaining distances that separate him from collision, it creates no inducement to vacate the border. Compellence, to be effective, can’t wait forever. Still, it has to wait a little; collision can’t be instantaneous. The compellent threat has to be put in motion to be credible, and then the victim must yield. [Too little time, and compliance becomes impossible; too much time, and compliance becomes unnecessary.] Thus compellence involves timing in a way that deterrence typically does not.

In addition to the question of “when,” compellence usually involves questions of where, what, and how much. “Do nothing” is simple, “Do something” ambiguous. “Stop where you are” is simple; “Go back” leads to “How far?” “Leave me alone” is simple; “Cooperate” is inexact and open-ended. A deterrent position—a status quo, in territory or in more figurative terms—can often be surveyed and noted; a compellent advance has to be projected as to destination, and the destination can be unclear in intent as well as in momentum and braking power. In a deterrent threat, the objective is often communicated by the very preparations that make the threat credible; the trip-wire often demarcates the forbidden territory. There is usually an inherent connection between what is threatened and what it is threatened about. Compellent threats tend to communicate only the general direction of compliance, and are
less likely to be self-limiting, less likely to communicate in the very design of the threat just
what, or how much, is demanded. . . .

Actually, any coercive threat requires corresponding assurances; [the object of a threat
is to give somebody a choice]. To say, “One more step and I shoot,” can be a deterrent threat
only if accompanied by the implicit assurance, “And if you stop I won’t.” Giving notice of
unconditional intent to shoot gives him no choice (unless by behaving as we wish him to
behave the opponent puts himself out of range, in which case the effective threat is, “Come
closer and my fire will kill you, stay back and it won’t”). What was said above about deter-
rent threats being typically less ambiguous in intent can be restated: the corresponding
assurances—the ones that, together with the threaten response, define the opponent’s
choice—are clearer than those that can usually be embodied in a compellent action. (Ordinary
blackmailers, not just nuclear, find the “assurances” troublesome when their threats are
compellent.)

They are, furthermore, confirmed and demonstrated over time; as long as he stays back,
and we don’t shoot, we fulfill the assurances and confirm them. The assurances that accom-
pany a compellent action—move back a mile and I won’t shoot (otherwise I shall) and I
won’t then try again for a second mile—are harder to demonstrate in advance, unless it be
through a long past record of abiding by one’s own verbal assurances.

Because in the West we deal mainly in deterrence, not compellence, and deterrent threats
tend to convey their assurances implicitly, we often forget that both sides of the choice, the
threatened penalty and the proffered avoidance or reward, need to be credible. The need for
assurances—not just verbal but fully credible—emerges clearly as part of “deterrence” in
discussions of surprise attack and “preemptive war.” An enemy belief that we are about to
attack anyway, not after he does but possibly before, merely raises his incentive to do what
we wanted to deter and to do it even more quickly. When we do engage in compellence, as
in the Cuban crisis or in punitive attacks on North Vietnam that are intended to make the
North Vietnamese government act affirmatively, the assurances are a critical part of the de-
definition of the compellent threat.

One may deliberately choose to be unclear and to keep the enemy guessing either to keep
his defenses less prepared or to enhance his anxiety. But if one wants not to leave him in
doubt about what will satisfy us, we have to find credible ways of communicating, and com-
municating both what we want and what we do not want. There is a tendency to emphasize
the communication of what we shall do if he misbehaves and to give too little emphasis to
communicating what behavior will satisfy us. Again, this is natural when deterrence is our
business, because the prohibited misbehavior is often approximately defined in the threatened
response; but when we must start something that then has to be stopped, as in compellent
actions, it is both harder and more important to know our aims and to communicate. It is par-
ticularly hard because the mere initiation of an energetic coercive campaign, designed for
compellence, disturbs the situation, leads to surprises, and provides opportunity and tempta-
tion to reexamine our aims and change them in mid course. Deterrence, if wholly successful,
can often afford to concentrate on the initiating events—what happens next if he misbehaves.
Compellence, to be successful, involves an action that must be brought to successful closure.
The payoff comes at the end, as does the disaster if the project fails.

The compellent action will have a time schedule of its own, and unless it is carefully
chosen it may not be reconcilable with the demands that are attached to it. . . . There will be
limits, probably, to how long the compellent action can be sustained without costing or risk-
ing too much, or exhausting itself or the opponent so that he has nothing left to lose. If it
cannot induce compliance within that time—and this depends on whether compliance is physically or administratively feasible within that time—it cannot accomplish anything (unless the objective was only an excuse for some act of conquest or punishment). The competent action has to be one that can be stopped or reversed when the enemy complies, or else there is no inducement.

If the opponent's compliance necessarily takes time—if it is sustained good behavior, cessation of an activity that he must not resume, evacuation of a place he must not reenter, payment of tribute over an extended period, or some constructive activity that takes time to accomplish—the compellent threat requires some commitment, pledge, or guarantee, or some hostage, or else must be susceptible of being resumed or repeated itself. Particularly in a crisis... there is strong incentive to get compliance quickly to limit the risk or damage. Just finding conditions that can be met on the demanding time schedule of a dangerous crisis is not easy. The ultimate demands, the objectives that the compellent threat is really aimed at, may have to be achieved indirectly, by taking pledges or hostages that can be used to coerce compliance after the pressure has been relieved. Of course, if some kind of surrender statement or acknowledgment of submission, some symbolic knuckling under, will itself achieve the object, verbal compliance may be enough. It is inherent in an intense crisis that the conditions for bringing it to a close have to be of a kind that can be met quickly; that is what we mean by an “intense crisis,” one that compresses risk, pain, or cost into a short span of time or that involves actions that cannot be sustained indefinitely. If we change our compellent threat from slow pressure to intense, we have to change our demands to make them fit the urgent timing of a crisis.

Notice that to deter continuance of something the opponent is already doing—harassment, overflight, blockade, occupation of some island or territory, electronic disturbance, subversive activity, holding prisoners, or whatever it may be—has some of the character of a compellent threat. This is especially so of the timing, of who has to take the initiative. In the more static case we want him to go on not doing something; in this more dynamic case we want him to change his behavior. The "when" problem arises in compelling him to stop, and the compellent action may have to be initiated, not held in waiting like the deterrent threat.

**DEFENSE AND DETERRENCE, OFFENSE AND COMPELLENCE**

The observation that deterrent threats are often passive, while compellent threats often have to be active, should not be pressed too far. Sometimes a deterrent threat cannot be made credible in advance, and the threat has to be made lively when the prohibited action is undertaken. This is where defense and deterrence may merge, forcible defense being undertaken in the hope, perhaps with the main purpose, of demonstrating by resistance that the conquest will be costly, even if successful, too costly to be worthwhile. . . . If the enemy act to be deterred is a once-for-all action, incapable of withdrawal, rather than progressive over time, any failure of deterrence is complete and final; there is no second chance. But if the aggressive move takes time, if the adversary did not believe he would meet resistance or did not appreciate how costly it would be, one can still hope to demonstrate that the threat is in force, after he begins. If he expected no opposition, encountering some may cause him to change his mind.

There is still a distinction here between forcible defense and defensive action intended to deter. If the object, and the only hope, is to resist successfully, so that the enemy cannot
succeed even if he tries, we can call it pure defense. If the object is to induce him not to proceed, by making his encroachment painful or costly, we can call it a "coercive" or "deterrent" defense. The language is clumsy but the distinction is valid. Resistance that might otherwise seem futile can be worthwhile if, though incapable of blocking aggression, it can nevertheless threaten to make the cost too high. This is "active" or "dynamic" deterrence, deterrence in which the threat is communicated by progressive fulfillment. At the other extreme is forcible defense with good prospect of blocking the opponent but little promise of hurting; this would be purely defensive.

Defensive action may even be undertaken with no serious hope of repelling or deterring enemy action but with a view to making a "successful" conquest costly enough to deter repetition by the same opponent or anyone else. This is of course the rationale for reprisals after the fact; they cannot undo the deed but can make the books show a net loss and reduce the incentive next time. . . . "Compellence" is more like "offense." Forcible offense is taking something, occupying a place, or disarming an enemy or a territory, by some direct action that the enemy is unable to block. "Compellence" is inducing his withdrawal, or his acquiescence, or his collaboration by an action that threatens to hurt, often one that could not forcibly accomplish its aim but that, nevertheless, can hurt enough to induce compliance. The forcible and the coercive are both present in a campaign that could reach its goal against resistance, and would be worth the cost, but whose cost is nevertheless high enough so that one hopes to induce compliance, or to deter resistance, by making evident the intent to proceed. Forcible action . . . is limited to what can be accomplished without enemy collaboration; compellent threats can try to induce more affirmative action, including the exercise of authority by an enemy to bring about the desired results.

War itself, then, can have deterrent or compellent intent, just as it can have defensive or offensive aims. A war in which both sides can hurt each other but neither can forcibly accomplish its purpose could be compellent on one side, deterrent on the other. Once an engagement starts, though, the difference between deterrence and compellence, like the difference between defense and offense, may disappear. There can be legal and moral reasons, as well as historical reasons, for recalling the status quo ante; but if territory is in dispute, the strategies for taking it, holding it, or recovering it may not much differ as between the side that originally possessed it and the side that coveted it, once the situation has become fluid. . . . The coercive aspect of warfare may be equally compellent on both sides, the only difference perhaps being that the demands of the defender, the one who originally possessed what is in dispute, may be clearly defined by the original boundaries, whereas the aggressor's demands may have no such obvious definition.

The Cuban crisis is a good illustration of the fluidity that sets in once passive deterrence has failed. The United States made verbal threats against the installation of weapons in Cuba but apparently some part of the threat was unclear or lacked credibility and it was transgressed. The threat lacked the automaticity that would make it fully credible, and without some automaticity it may not be clear to either side just where the threshold is. Nor was it physically easy to begin moderate resistance after the Russians had crossed the line, and to increase the resistance progressively to show that the United States meant it. By the time the President determined to resist, he was no longer in a deterrent position and had to embark on the more complicated business of compellence. The Russian missiles could sit waiting, and so could Cuban defense forces; the next overt act was up to the President. The problem was to prove to the Russians that a potentially dangerous action was forthcoming, without any confidence that verbal threats would be persuasive and without any desire to initiate
some irreversible process just to prove, to everybody’s grief, that the United States meant what it said.

The problem was to find some action that would communicate the threat, an action that would promise damage if the Russians did not comply but minimum damage if they complied quickly enough, and an action that involved enough momentum or commitment to put the next move clearly up to the Russians. Any overt act against a well-defended island would be abrupt and dramatic; various alternatives were apparently considered, and in the end an action was devised that had many of the virtues of static deterrence. A blockade was thrown around the island, a blockade that by itself could not make the missiles go away. The blockade did, however, threaten a minor military confrontation with major diplomatic stakes—an encounter between American naval vessels and Soviet merchant ships bound for Cuba. Once in place, the Navy was in a position to wait; it was up to the Russians to decide whether to continue. If Soviet ships had been beyond recall, the blockade would have been a preparation for inevitable engagement; with modern communications the ships were not beyond recall, and the Russians were given the last clear chance to turn aside. Physically the Navy could have avoided an encounter; diplomatically, the declaration of quarantine and the dispatch of the Navy meant that American evasion of the encounter was virtually out of the question. For the Russians, the diplomatic cost of turning freighters around, or even letting one be examined, proved not to be prohibitive.

Thus an initial deterrent threat failed, a compellent threat was called for, and by good fortune one could be found that had some of the static qualities of a deterrent threat.

There is another characteristic of compellent threats, arising in the need for affirmative action, that often distinguishes them from deterrent threats. It is that the very act of compliance—of doing what is demanded—is more conspicuously compliant, more recognizable as submission under duress, than when an act is merely withheld in the face of a deterrent threat. Compliance is likely to be less casual, less capable of being rationalized as something that one was going to do anyhow. . . . Indeed, if a deterrent threat is created before the proscribed act is even contemplated, there need never be an explicit decision not to transgress, just an absence of any temptation to do the thing prohibited.

The Russians cannot, though, claim that they were on the point of removing their missiles from Cuba anyway, and that the President’s television broadcast, the naval quarantine and threats of more violent action, had no effect. . . .

If the object is actually to impose humiliation, to force a showdown and to get an acknowledgment of submission, then the “challenge” that is often embodied in an active compellent threat is something to be exploited. President Kennedy undoubtedly wanted some conspicuous compliance by the Soviet Union during the Cuban missile crisis, if only to make clear to the Russians themselves that there were risks in testing how much the American Government would absorb such ventures. . . .

“CONNECTEDNESS” IN COMPETENT THREATS

As mentioned earlier, a deterrent threat usually enjoys some connectedness between the proscribed action and the threatened response. The connection is sometimes a physical one, as when troops are put in Berlin to defend Berlin. Compellent actions often have a less well-defined connectedness; and the question arises whether they ought to be connected at all. If the object is to harass, to blockade, to scare or to inflict pain or damage until an adversary complies, why cannot the connection be made verbally? If the Russians want Pan-American
Airways to stop using the air corridor to Berlin, why can they not harass the airline on its Pacific routes, announcing that harassment will continue until the airline stops flying to Berlin? When the Russians put missiles in Cuba, why cannot the President quarantine Vladivostok, stopping Soviet ships outside, say, a twelve-mile limit, or perhaps denying them access to the Suez or Panama Canal? And if the Russians had wanted to counter the President’s quarantine of Cuba, why could they not blockade Norway?

A hasty answer may be that it just is not done, or is not “justified,” as though connectedness implied justice, or as though justice were required for effectiveness. Surely that is part of the answer; there is a legalistic or diplomatic, perhaps a casuistic, propensity to keep things connected, to keep the threat and the demand in the same currency to do what seems reasonable. But why be reasonable, if results are what one wants? Habit, tradition, or some psychological compulsion may explain this connectedness, but it has to be asked whether they make it wise.

There are undoubtedly some good reasons for designing a compellent campaign that is connected with the compliance desired. One is that it helps to communicate the threat itself; it creates less uncertainty about what is demanded and what pressure will be kept up until the demands are complied with and then relaxed once they are. Actions not only speak louder than words on many occasions, but like words they can speak clearly or confusingly. To the extent that actions speak, it helps if they reinforce the message rather than confuse it.

Second, if the object is to induce compliance and not to start a spiral of reprisals and counteractions, it is helpful to show the limits to what one is demanding, and this can often be best shown by designing a campaign that distinguishes what is demanded from all the other objectives that one might have been seeking but is not. To harass aircraft in the Berlin air corridor communicates that polar flights are not at issue; to harass polar flights while saying that it is punishment for flying in the Berlin corridor does not so persuasively communicate that the harassment will stop when the Berlin flights stop, or that the Russians will not think of a few other favors they would like from the airline before they call off their campaign. Most of the problems of defining the threat and the demands that go with it, of offering assurance about what is not demanded and of promising cessation once compliance is forthcoming, are aggravated if there is no connection between the compellent action (or the threat of it) and the issue being bargained over.

The same question can arise with deterrent threats; sometimes they lack connectedness. To threaten the Chinese mainland in the event of an overland attack on India has a minimum of connectedness. If the threatened response is massive enough, though, it may seem to comprise or to include the local area and not merely to depart from it. But it often lacks some of the credibility, through automatic involvement, that can be achieved by connecting the response physically to the provocation itself. Contingent actions—not actions initiated to induce compliance, but actions threatened against potential provocation—often need the credibility that connectedness can give them.

Connectedness in fact provides something of a scheme for classifying compellent threats and actions. The ideal compellent action would be one that, once initiated, causes minimal harm if compliance is forthcoming and great harm if compliance is not forthcoming, is consistent with the time schedule of feasible compliance, is beyond recall once initiated, and cannot be stopped by the party that started it but automatically stops upon compliance, with all this fully understood by the adversary. Only he can avert the consequences; he can do it only by complying; and compliance automatically precludes them. His is then the “last clear chance” to avert the harm or catastrophe; and it would not even matter which of the two
most feared the consequences as long as the adversary knew that only he, by complying, could avert them. (Of course, whatever is demanded of him must be less unattractive to him than the threatened consequences, and the manner of threatened compliance must not entail costs in prestige, reputation, or self-respect that outweigh the threat.)

It is hard to find significant international events that have this perfectionist quality. . . .

COMPELLENCE AND BRINKMANSHIP

Another important distinction is between compellent actions that inflict steady pressure over time, with cumulative pain or damage to the adversary (and perhaps to oneself), and actions that impose risk rather than damage. Turning off the water supply at Guantanamo creates a finite rate of privation over time. Buzzing an airplane in the Berlin corridor does no harm unless the planes collide; they probably will not collide but they may and if they do the result is sudden, dramatic, irreversible, and grave enough to make even a small probability a serious one.

The creation of risk—usually a shared risk—is the technique of compellence that probably best deserves the name of “brinkmanship.” It is a competition in risk-taking. It involves setting afoot an activity that may get out of hand, initiating a process that carries some risk of unintended disaster. The risk is intended, but not the disaster. One cannot initiate certain disaster as a profitable way of putting compellent pressure on someone, but one can initiate a moderate risk of mutual disaster if the other party’s compliance is feasible within a short enough period to keep the cumulative risk within tolerable bounds. “Rocking the boat” is a good example. If I say, “Row, or I’ll tip the boat over and drown us both,” you’ll not believe me. I cannot actually tip the boat over to make you row. But if I start rocking the boat so that it may tip over—not because I want it to but because I do not completely control things once I start rocking the boat—you’ll be more impressed. I have to be willing to take the risk; then I still have to win the war of nerves, unless I can arrange it so that only you can steady the boat by rowing where I want you to. But it does lend itself to compellence, because one may be able to create a coercive risk of grave consequences where he could not profitably take a deliberate step to bring about those consequences, or even credibly threaten that he would. . . .

NOTE

1. The critical role of “assurances” in completing [the structure of a threat], in making the threatened consequences persuasively conditional on behavior so that the victim is offered a choice, shows up in the offers of amnesty, safe passage, or forgiveness that must often be made credible in inducing the surrender of rebels or the capitulation of strikers or protesters. Even libraries and internal revenue agencies depend on parallel offers of forgiveness when they embark on campaigns to coerce the return of books or payment of back taxes. In my personal life I have sometimes relied, like King Lear, on the vague threat that my wrath will be aroused (with who knows what awful consequences) if good behavior is not forthcoming, making a tentative impression on one child, only to have the threat utterly nullified by another’s pointing that “Daddy’s mad already.”