General deterrence and the balance of power

Lawrence Freedman*

The advent of nuclear weapons is clearly the single most important development which post-war balance of power thinking has had to accommodate. Although there have been earlier attempts to explore this relationship,1 an examination of the current debate on deterrence should throw some light on the issue of the continuing relevance of the balance of power concept.

It often appears in fact as if two separate debates on deterrence are taking place side by side. One is the familiar policy debate, preoccupied with the sustenance of superpower deterrence, either in the narrow version which is largely concerned with preventing a surprise attack against the continental United States, or the extended version which is required to 'couple' the United States to its allies. Most of the various positions in this policy debate were established some years ago and concern such matters as the virtues and vices of 'first-use threats' and the balance between conventional and nuclear forces. It has a ritualistic quality. Sometimes it serves as a coded means of talking about the distinctive security interests within the alliance without sounding too blunt; more often it just provides the best theory available to guide the developments of force plans and doctrine.

The categories with which strategy is discussed in the West have been forged through this debate. It has shaped language and thought to the point where it is difficult for the NATO establishment to talk about strategy in any other way. The debate is political rather than analytical, in that few of the propositions put forward during its course can be properly tested until East-West relations reach a level of crisis that has been successfully avoided up to now.

The second debate is analytical rather than political. The academic debate began with an effort to make sense of the superpower relationship. With the steady loss of American nuclear superiority, what was called 'deterrence' was being practised in less-than- optimum circumstances. Academics began to explore how it might none the less still operate. By and large this exploration took the form of abstract theorizing, although unusually this quality by no means diminished its impact. Robert Jervis has described it as 'probably the most influential school of thought in the American study of international relations', although, as Jervis also notes, it was derivative of Realism.2

As impatience with the abstract nature of deterrence theory grew a more empirical approach came to be adopted. Following the lead taken in the path-breaking study by Alexander George and Richard Smoke,3 a series of investigations of deterrence in practice, on occasion involving the superpowers but as often as not involving

* A number of paragraphs in this article, especially in its second half, resemble those in my paper on 'Extended Nuclear Deterrence', presented to the conference of the International Institute for Strategic Studies in Brighton in September 1988 and to be published in the conference proceedings. The two pieces represent the elaboration of the same argument, but to two different audiences.
different countries and even different eras, have appeared. This literature concentrates on the problem of manipulating another's potentially aggressive behaviour through threats. This problem is in principle similar to that at the heart of the policy debate. Moreover the academic investigations of deterrence in practice are by no means innocent of the implications for western policy, given that a number of the cases studied bear a resemblance to the standard NATO nightmare and illuminate a series of flaws in deterrence as an approach to crisis management. However, for precisely the same reason that the more abstract deterrence theorists considered their abstraction warranted because of the sheer novelty of the situation, there is inevitably a sense that the conclusions derived from case studies must inevitably be limited because we have not yet experienced that ultimate test which is at the heart of the policy debate. So while it would be unfair to suggest that the thorough research has lacked influence it has not, as it were, 'deterred' the NATO establishment from continuing to phrase their strategic pronouncements in terms of deterrence.

It is not my intention in this article to challenge the particular conclusions on deterrence as analytical theory and strategic practice which have been developed in recent years by a number of scholars. They contain much insight and good sense. My real concern is with the analytical focus on deterrence as an immediate problem of crisis management rather than as a more general feature of international relations. This tendency is also exhibited in some contributions to the policy debate. This article explores the concept of general deterrence, which is the closest deterrence theory comes to traditional balance of power notions. Although this concept was first offered over a decade ago, it has not been given the attention that it deserves by the strategic studies community.

The concept of deterrence

The early use of the term deterrence in twentieth century strategic thinking followed closely the Latin root—deterre or to frighten from or away. The word has seemed most appropriate when being used to convey the idea of scaring off a potential aggressor through threats of consequential pain. However, once a more precise formulation is attempted then the concept moves away from its root. Consider for example the definition employed by George and Smoke:

Deterrence is simply the persuasion of one's opponent that the costs and/or risks of a given course of action he might take outweigh its benefits.

Such a formulation covers a wide range of strategic relationships. The element of fear is not so prominent: calculations of costs and risks are made by planners all the time without any surge of adrenalin or a quickening of the pulse. If the key question is whether the costs and/or risks outweigh the benefits, this can be a modest judgement concerning 'value-for-money' and can be influenced as much by an assessment of the extent of the likely benefits as of the likely costs. Indeed a major distinction in deterrence literature takes the possibility of encouraging the opponent to expect only limited gains as a plausible and in many ways more satisfactory alternative to encouraging a prospect of severe pain, especially in conditions when the opponent can threaten severe pain in return. Others argue that the concept of deterrence is incomplete without as much attention to inducements and reassurance as to threats. The definition makes it clear the idea is to dissuade the opponent from initiating action rather than to compel him to do—or undo—something against his will, again reflecting a distinction developed in the literature. However, it is by no means clear that the 'something' in question threatens the deterrer directly. This raises another
familiar problem: what range of interests can be protected through deterrence? Lastly the definition acknowledges that the success of deterrence depends on the opponent being persuaded. No matter how sincere the deterrer might be in his conditional threats, if the opponent does not take these threats seriously then deterrence will fail.

This leads on to a central theme in both the policy and the academic debates. If deterrence is in the eye of the beholder then the opponent may simply misapprehend the message that he is being sent and fail to act accordingly. The problem with designing deterrence strategies has therefore been to find ways of ensuring that the opponent receives the threat, relates it to his proposed course of action and decides as a result not to go ahead as planned. The use in the definition from George and Smoke of the phrase ‘costs and/or risks’ recognizes that the opponent need not be convinced that the costs will definitely be imposed, only that there is a significant probability of this being so.

For purposes of analysis this peculiar quality of deterrence, with the opponent being persuaded not to do something, makes it very difficult to know whether in practice a deterrence relationship is in being. If the opponent is inactive this may be because he has no inclination to act, or, if he has been persuaded not to act, this may be for reasons quite separate from the deterrer or from the particular character of deterrent threats. This is often discussed as a problem for the deterrer. Is he wasting his time by making an effort to deter something that cannot be deterred or does not need deterring? How can he make his threats sufficiently credible to penetrate the mind-set of his opponent? Does this credibility depend on really being prepared to carry out the threat or merely conveying a sufficient probability that he just might?

But it is also a problem for the deterred. Is he missing an opportunity because of mythical fears about the possible consequences? The condition of paranoia, which is much discussed in the deterrence literature, is an obvious example of being influenced by fear of another which has little basis in reality. A deterrer can remain innocent of his influence on an opponent’s calculations without the opponent losing his grip on reality. It is possible, even normal, to be persuaded against a particular course of action by the thought of how the target might respond. Prudence might dictate caution without the potential target being aware that he had ever been at risk. A would-be aggressor may thus be effectively deterred by an accurate assessment of the likely form of his potential victim’s response without the victim having to do very much.

In the policy debate the phrase ‘self-deterrence’ is sometimes used to denote an unwillingness to take necessary initiatives as a result of a self-induced fear of the consequences. But all deterrence is self-deterrence in that it ultimately depends on the calculations made by the deterred, whatever the quality of the threats being made by the deterrer. So while much of the discussion of deterrence revolves around the problem of adopting it as a strategy, analytically it is important to recognize that it is as interesting to examine it from the perspective of the deterred as from that of the deterrer. Moreover deterrence can seem far less problematic when we start from the point of view of the deterred. Once certain courses of action have been attempted this conclusion may be institutionalized. It requires little further deliberation. Can the concept still be useful if it is used to describe a continuing feature of a state’s response to the international environment? I would argue that it can because the conclusion is still conditional. A change in the particular relationship could activate those strategic options that have been ruled out for the moment. More importantly, in practice this is how superpower deterrence works. Much of the debate on deterrence can seem to be beside the point because it focuses on a situation that has yet to arise rather than attempting to explain the one that exists already.
Immediate and general deterrence

To help us explore this further we can take advantage of Patrick Morgan’s distinction between immediate and general deterrence.

*Immediate deterrence* concerns the relationship between opposing states where at least one side is seriously considering an attack while the other is mounting a threat of retaliation in order to prevent it. *General deterrence* relates to opponents who maintain armed forces to regulate their relationship even though neither is anywhere near mounting an attack.10

Morgan notes that while general deterrence is more typical, the theory ‘has been developed almost exclusively by hypothesizing an abstract world of immediate deterrence’.11 He suggests that it would be wiser to start with general deterrence for: severe crises usually emerge out of general deterrence (general hostility) relationships, and they must be handled by governments that cannot simply step into the abstract, rational world of classic deterrence theory but must remain somewhat mired in the personal quirks, ideological pressures, and internal-societal political pressures of non-crisis times. Thus the proper direction for our analytical perspective has been reversed. Deterrence theory should begin with what we know about individual and governmental behavior in more normal times, trimming and adjusting that framework to fit the unusual circumstances of a severe confrontation when immediate deterrence is needed.12

Yet at the end of his book Morgan notes that

General deterrence is . . . something of a residual category and has not been given detailed attention in this book. . . . [It] deserves much more attention, something left for a later work.

As far as I am aware this later work remains to be written. Theorists of deterrence have echoed Morgan’s own complaint that general deterrence is too ‘amorphous’.13 Immediate deterrence has been considered more important and also more researchable. Thus a study of general deterrence may tell us little about the conditions in which immediate deterrence can function, while it may be difficult to ascertain in practice whether general deterrence is actually operating at all.14 Lebow and Stein note that those interested in exploring general deterrence are deprived of an ‘evidentiary trail’ and lack clear criteria for case selection:

[T]he success of general deterrence can be inferred only through counterfactual argument rather than evidence and is often subject to intense controversy. For this reason, analysts of deterrence generally restrict their selection to cases of immediate deterrence.15

This is true of only one set of analysts of deterrence. Those who focus on East-West relations are essentially considering general deterrence, and on the conduct of these relations there is ample evidence.16 Much of the policy debate consists of an interrogation of the past to suggest how a relatively satisfactory situation might be sustained. Because there are numerous factors making for the absence of war this interrogation is less than conclusive. It is none the less more fruitful than many attempts to anticipate circumstances in the form of ‘scenarios’ which bring the two alliances to the point of confrontation in order to assess how immediate deterrence might then be practised.
Morgan describes general deterrence in terms of a set of decision makers stopping short of seriously considering an attack in response to a rather vague threat by another government, but with the clear implication that without that vague threat they might have resorted to force. But stopping short is not a state of suspended animation. As Morgan notes, general deterrence can be a stance of many years' duration. The decision makers will not simply pick up where they left off should the activity which constituted the original 'vague threat' cease some time later, but will be obliged to reconsider their policy in terms of all that has transpired in the intervening period. 'Stopping short' becomes institutionalized, so that it would take a deliberate decision to once again consider the resort to force.

The discussion of the distinction between immediate and general deterrence by Lebow and Stein reinforces this point. Immediate deterrence they characterize in terms of 'a strategy of conflict management', involving a conscious effort by leaders to 'manipulate the risk of war in order to influence the behaviour of an adversary'. By contrast general deterrence is 'an expression of existing power relationships . . . quite independent of efforts by a state to define and publicize a commitment, and to threaten war in its defense'. Thus the key feature of general deterrence is that it is not a strategy—a deliberate effort 'to dissuade a challenger from resorting to force'.

Morgan, by contrast, does see general deterrence as involving a strategy, but it is a different sort of strategy from that involved with immediate deterrence.

At the level of general deterrence, then, arms and warnings are a contribution to the broad context of international politics, to the system within which the state seeks its security. The idea is to manage the context so that for an opponent it will appear basically unattractive to resort to force.

He notes that this sort of notion is 'often what analysts and statesmen have had in mind when referring to the operations of a balance of power'. This takes us back to Lebow's and Stein's idea that general deterrence is something that resides in power relationships rather than requiring active pursuit.

Governments do make statements and take action, including the procurement of weapons systems, in the name of deterrence, so it is hard to say that a condition of general deterrence can exist independent of strategy. A more useful way of describing the difference might be in the degree of strategic engagement between deterrer and deterred. Immediate deterrence involves an active effort to deter in the course of a crisis when the efficiency of any threats will soon be revealed in adversary behaviour. General deterrence is altogether more relaxed, requiring merely the conveyance of a sense of risk to a potential adversary to ensure that active hostilities are never seriously considered. The lack of direct engagement means that deterrence depends largely on the deterred's assessment of his strategic environment. If hostilities are not seriously considered that reflects the strategic perspective of the deterred.

There is an urgency to immediate deterrence. It is part of the atmosphere of crisis and so can be recognized in the government papers of the late 1930s, or the late 1940s. It can be detected in the 1950 document NSC-68, which argued the need for massive US rearmament in the face of the Communist challenge, and identified 1954 as a 'year of maximum peril'. When 1954 came and the peril appeared manageable there was a deliberate shift to a preparation for the 'long haul'. That was the point of transition from immediate to general deterrence. Only in the early 1960s, with the Berlin and Cuban crises, was there a serious hint of a return to immediate deterrence. Since then East-West relations have been characterized by general deterrence, and if we confine ourselves to the most important criterion of all—the avoidance of total war—the result has been satisfactory.
Although Lebow has been to the fore in developing a critique of immediate deterrence by means of case studies, he acknowledges the value of general deterrence in attempts to explain the cold war. He claims that it has been more successful than often realized, and certainly more so than immediate deterrence:

The strategy of deterrence has had largely pernicious effects. General deterrence, by contrast, has had a moderating influence on superpower behaviour. It rests on the realization in Moscow and Washington that there is nothing that can be done to protect either country from the near-certain annihilation that would result from a superpower nuclear war.21

This statement acknowledges that in general deterrence the critical strategic decision is taken by the deterred when he chooses not to push in a particular direction out of an awareness of the potential catastrophe which could result. In this way, the success of general deterrence is tied up with the unambiguous horror of nuclear weapons.

Lebow also raises a number of problems, incidentally accepting that general statements cannot be considered simply as a condition but must contain a strategic element. He considers deterrence strategies as tending to encourage an exaggerated view of the importance of demonstrations of ‘resolve’ in the face of challenges that would otherwise be recognized as minor, participation in an arms race, and a degree of sustained antagonism in a political relationship. Furthermore, there is no guarantee that it will always succeed ‘in keeping leaders from behaving in ways that could lead to war’.22 However these concerns assume that the strategy involved in general deterrence is broadly the same as that involved in immediate deterrence. As general deterrence lacks the close engagement between the two adversaries peculiar to immediate deterrence, the risks of the effects Lebow considers ‘pernicious’ may be correspondingly lower.

Morgan sees the challenge to general deterrence coming less from the logic of a particular strategy required to sustain it, than from its possible inability to cope with turbulence within the international system of the sort that has produced wars in the past, and in particular political changes which create dissatisfaction with the established international order. General deterrence becomes frail at times of revolutionary change or when the international order appears illegitimate in the eyes of rising states. History is therefore full of cases in which general deterrence, in the guise of a balance of power, has broken down. However, it has yet to do so when the balance is based on nuclear power, so the relevance of the pre-nuclear cases remains unclear. The issue is important: a move from general to immediate deterrence means that one form of deterrence has already broken down. General deterrence is practised in order to avoid having to practise immediate deterrence. If we can manage general deterrence then we have less to fear from the inadequacies of immediate deterrence.

**Practical implications**

For NATO this may be just as well, for it is by no means clear that the doctrine of flexible response could survive a close engagement with the Warsaw Pact. Of course many in the NATO establishment believe that this doctrine offers useful guidance in the event of sustained Warsaw Pact aggression. But in practice ‘flexible response’ is geared to peacetime. It offers a form of words that can accommodate the great variety of strategic perspectives found within the alliance. It works in peacetime precisely because it does not prescribe for actual hostilities. The language of flexibility and options avoids confronting the dilemmas of coalition warfare. It postpones the most difficult of all strategic decisions to a hypothetical time when they could be
postponed no longer but when they could also at least be made by reference to the developing political and military context.

Once nuclear deterrence is discussed as an immediate challenge then it soon appears difficult and even hopeless. This is more than confirmed by much of the analytical literature which draws on the history of international relations and points to the difficulty of manipulating threats to influence an adversary’s behaviour.23 What is put into these threats is rarely what is received. Examples can be found of misunderstandings and confusion which either fail to deter someone who needs deterring, or else aggravates a crisis that might otherwise have been managed peacefully. Military signals in particular are often notoriously ambiguous and the problems of interpretation grow in the psychological intensity of crisis. Even if threats can be well constructed and are perfectly understood when a conflict reaches crisis point, they are still only one part of the equation: the interests at stake, the underlying political trends, the attitude of allies and so on must also be considered. It is only one variable among many and not necessarily the most important.

In the policy debate the problem of immediate nuclear deterrence is addressed by speculating on the specifics of a future East–West confrontation and wondering how on earth the West will cope. The starting point is the failure of deterrence and the fear that unless one has something sensible to do should it fail then it is no more than bluff. Eventually it will be exposed as such and so deterrence will fail. But the success of general deterrence need not depend on a speculative assessment of performance against the far more demanding criteria of immediate deterrence.

The possibility of catastrophic miscalculation by either side casts doubt on the validity of a strategy of immediate deterrence in an intense crisis. However, this same possibility if anything reinforces general deterrence in that it reinforces a disinclination to war. Both types work, if at all, through the adversary recognizing an unacceptable risk of unacceptable costs for inadequate gain. In the immediate version this must be a complex calculation, while in the general version all that is needed is a vague feeling that certain options are just not worth pursuing. Nuclear war, even if the worst excesses of city busting and ‘nuclear winter’ can be avoided, will be so horrific that only the most extreme political pressure would make it worth tolerating the slightest risk of this coming to pass. So long as there is a measurable risk that a confrontation could end up with nuclear detonations there is every incentive to avoid the confrontation.

This residual risk saved nuclear deterrence once the preferred approach based on a clear-cut superiority no longer seemed plausible. There is no strategic formula to prescribe what is necessary to sustain this risk. It is about tolerating processes of great potential irrationality which create the chance of nuclear weapons being detonated. Circumstances discussed in the literature include a desperate and intense conventional fight in which passions have been raised, rash decisions become more likely, and the overall situation slips out of control, and/or some catastrophic series of accidents which lead to nuclear weapons being launched where better information and control would have made this wholly impossible.24

Recent ideas about ‘existential deterrence’, according to which the mere possibility that nuclear weapons could be used creates a sufficient deterrence effect,25 have an intuitive strength largely because we are aware of conflicts in which reckless and terrible deeds have been done by people or states who would not have believed themselves capable of such things beforehand. The uncertainties and pressures of war cast doubt on attempts to construct doctrines for nuclear use and so in turn make it extremely difficult to rely on such doctrines as a basis for reliable deterrent threats. By the same token they give sufficient credibility to deterrence based on the prospect
of a confrontation getting out of control, and disaster through inadvertence. General deterrence works through these assumptions becoming internalized by the deterred.

**Role of the deterrer**

What role does this leave for the deterrer? Is it the case that so long as a degree of military preparedness is maintained, general deterrence follows naturally? I would suggest that the key requirement for the deterrer is less to give credibility to the threat, than to clarify the circumstances in which the threat would be activated. The superpowers cannot be expected to run the risks of nuclear war in support of peripheral interests. This is why states of peripheral interest to their natural superpower supporter, but of core interest to their superpower adversary, may become less than convinced that deterrence is working on their behalf.

A central issue in US policy-making from the start of the cold war has been over how far the concept of vital interests can be extended to ensure a link with nuclear deterrence. The experience of SEATO and CENTO suggests that merely to declare alliance is not enough, and that policy-makers make distinctions about the risks they are prepared to run, and the resources they are prepared to devote, to different regions of the world. Western Europe, Japan and South Korea rate higher than other regions. This is to some extent for historical and institutional reasons as much as the result of an analysis of objective interests. The argument became muddied by such notions as the 'interdependence of commitments', which suggested that it was vital to meet any communist challenge to the friends of the United States (neither of which were necessarily precisely defined) in order that there be no doubt as to American resolve. Although this argument occasionally resurfaces, it lost much of its force as a result of the Vietnam war. There is an awareness of the limits to which one can extend deterrence. These limits are less a function of the strategic balance itself (especially since the United States lost any semblance of superiority) and more of a definition of vital interests.

The military postures adopted by the major powers are hardly irrelevant, but they are still only one of a number of factors which shape the current security system. At times of immediate deterrence these postures become much more important. Then they have to be fashioned carefully with regard to political positions. The hypothetical problem at the heart of the debate on giving credibility to deterrent threats has not yet become critical in practice because the political context has not obliged political leaders to address the stark choices posed.

The extension of general deterrence may require little more than acknowledging the possibility of becoming caught up in a total war as a result of commitments made to another country. If that possibility is significant then by definition the security of that country has become a vital interest. Hence the focus within the alliance of creating possibilities for a local conflict to escalate to a superpower level. The possibility of military instability creates the conditions for political stability.

This brings into question the familiar distinction between extended and central deterrence. While they share a number of fundamental features, especially the belief that an adversary's aggressive inclination can be controlled through threats, the underlying strategic relationships are often characterized as being quite different and possibly even contradictory.26

Extended deterrence appears much more difficult than central deterrence. The American homeland is only really vulnerable to nuclear attack, and to deter such an attack all that is needed is the threat of nuclear retaliation in kind. This is credible because attacks on sovereign territory—the most vital of all vital interests—are con-
sidered to provide sufficient motivation to order a nuclear riposte. With extended deterrence on the other hand it is necessary to be prepared to launch a nuclear attack in response to a conventional attack against third parties. This strains credibility.

Yet the possibility of a threat, nuclear or otherwise, to the United States only arises in the first place because of its overseas commitments. If the United States were truly isolationist, it need not clash with any other power, almost by definition. Once the United States decides not to be indifferent to what is going on in the rest of the world it is obliged to accept the risks that come with engagement. An isolationist United States could also want a nuclear arsenal to ward off any hostile powers; this potential was seen by many adherents of the new weapons to be their greatest virtue at the start of the atomic age. However, it is highly unlikely that a nuclear arsenal developed for this purpose would be assessed against the same sort of exacting standards that have in practice been adopted by the United States over the past decades.

It may therefore be the case that there is no serious requirement to deter a direct challenge to the United States. If the United States is under a threat this is because it is intimately involved in the affairs of other regions. Capabilities to attack the United States might exist, but it is only the extension of American interests that creates potential conflicts with other powers.

Central deterrence therefore became necessary because of alliance commitments. The greater the vulnerability of the United States the less likely it has appeared that it would take risks on behalf of allies. The problem of vulnerability would be most acute if the United States could be disarmed in a surprise attack. Despite preoccupations with the survivability of land-based ICBMs and such like, this fear has, to all intents and purposes, been abated. For exactly the same reasons, the United States cannot protect itself against Soviet retaliation should it initiate nuclear hostilities. This is why the problem of extended deterrence has now narrowed down to the credibility of a first use threat.

Thus rather than extended deterrence being a distinct type of deterrence it may be better understood as an attempt to achieve the maximum deterrent effect from a particular strategic relationship. In this sense one might as usefully describe central deterrence as contracted deterrence to indicate shrunken political objectives, or previously ‘vital’ interests being consigned to the periphery.

The strategy of general deterrence is therefore bound up with the process of defining vital interests; ambiguity on this score, brought about through the process of political change, is therefore rightly recognized to pose a challenge.

**Stability or instability?**

To what extent does general deterrence tend to instability? In a recent article John Mueller has outlined a distinction between two forms of stability that go in parallel with that between immediate and general deterrence. The distinction is between crisis stability and general stability. He notes how the literature on defence policy has been preoccupied with the former, and looks to both sides enjoying confidence in their ability to survive a surprise attack, thereby removing any incentive to initiate a war. General stability derives from broader needs, desires and concerns.

It prevails when two powers, taking all potential benefits, costs and risks into account, generally prefer peace to war—in the extreme, even to a victorious war—whether crisis stability exists or not.27

There are a number of problems with the notion of general stability which Mueller
does not fully explore. The important point about concepts such as crisis stability and
deterrence is that they refer to relations between antagonistic states. Mueller cites as
an example of general stability relations between Canada and the United States. This
makes the point that overwhelming military advantage does not necessarily result in
war, but it is hard to be excited by this observation when the two countries have so
little to fight about. The concept would be more useful if narrowed down to refer to
relations between states or groups of states who consider each other to be potential
enemies in their military planning but have reached a degree of mutual political
understanding that prevents this potential enmity being realized. Thus general
stability—and for that matter general deterrence—no longer applies once it has
become so successful that it dissolves into an essentially non-antagonistic relation-
ship. This may be reflected in a dissolution, or at least erosion, of rival alliances. It
will no longer be operating when an alliance loses its broad political identity, either
through internal divisions or the lines between the two antagonists becoming blurred
at the edges.

The other possibility is a move into immediate deterrence, at which point questions
of crisis stability become pertinent. As discussed earlier, this will most probably
result from political change. Even without an enquiry into bipolar versus multipolar
systems or the sorts of regimes that are likely to feel more comfortable with the estab-
lished order, it is clear that some political arrangements will be more stable than
others and will generate fewer conflicts.

If a relationship begins to develop instability and tends towards breakdown, when
war is still distant but just about conceivable, then a concern with capabilities for
immediate deterrence may develop. Prior to that point the policy problem is to
maintain options should immediate deterrence become necessary when these options
are being developed in the political climate created by years of general deterrence.
Capabilities may well deteriorate, with the only possible consolation a comparable
deterioration in those of the adversary. It is only at this point that competitive
military procurement, that is, arms racing, may come to the fore. During times of
general deterrence, each side's force structure will be shaped by a variety of
economic, technical, cultural and political factors as well as an assessment of the
other's force structure. In times of immediate deterrence states become more
concerned about military efficiency and combat readiness. In the East-West
relationship, we have barely moved close to the serious practice of immediate
deterrence, but there have been points when it was possible that the process of
transition from general to immediate was starting. In practice, as would be suggested
from this analysis, the major aggravating factors in East-West relations, and a spur
to increased defence budgets, have been political initiatives and/or military inter-
ventions which appear to pose a challenge to the vested interests of the other
side—Berlin, Vietnam, Afghanistan. Thus a condition of instability within general
deterrence might be one in which a deterioration in political relations was aggravated
by a surge of military activity. If front-line forces decline too far in relation to
adversary capabilities then there is a risk of even a mild crisis being aggravated by a
frantic rush to get forces in place.

To conclude: general deterrence has been described as an institutionalized
perception by a state or group of states that, despite continuing antagonism, it should
not expect to be able to resolve its disputes with another state or group of states by
military means. This will reflect an understanding of the vital interests of the deter-
ror which the deterred accepts must be respected. The longer this condition lasts the more
stable it is likely to become, with a growing tolerance of disparities in military capa-
bilities and also of political change. General stability can thus be described as a
deterrence relationship which tends towards the dissolution of antagonism; general instability suggests a tendency towards crisis, and the consequent practice of immediate deterrence.

References and notes

5. George and Smoke, op. cit., p. 11.
7. This is essentially the position taken by Lebow and Stein, op. cit.
8. The distinction between deterrence and compellence, i.e. between ‘inducing inaction and making someone perform’, has been most fully elaborated by Thomas Schelling, *Arms and Influence* (New Haven, Conn., 1966), p. 175, pp. 69ff.
11. Ibid., p. 29.
12. Ibid., p. 30.
13. Ibid., p. 204. He reports that he only included the concept in a second draft of his book after colleagues had complained that his total concentration on immediate deterrence neglected the constancy in the US–Soviet relationship and the extent to which deterrence was required to work not only at the point when a Soviet attack was under active consideration but also to keep them from ever getting to this point.
17. Morgan, op. cit., pp. 40–3. He outlines the characteristics of general deterrence as: (1) Relations between opponents are such that leaders in at least one would consider resorting to force if the opportunity arose. (2) The other side, precisely because it believes the opponent would be willing to consider resorting to force, maintains forces of its own and offers warnings to respond in kind to attempts to use force contrary to its interests. (3) The decision makers at whom the general deterrent threat is aimed do not go beyond preliminary consideration of resorting to force because of the expectation that such a policy would result in a corresponding resort to force of some sort by leaders of the opposing state.
19. Morgan, op. cit., p. 43.
20. Ibid., p. 42.
22. Ibid., p. 188.
23. R. Jervis, R. N. Lebow, J. G. Stein, *Psychology and Deterrence* (Baltimore, Md., 1985) along with the works by these authors cited earlier (n. 2 and n. 4) and those by George and Smoke (n. 3) and Huth and Russett (n. 14).
General deterrence and the balance of power


28. Mueller notes that 'the vast majority of wars that never take place are caused by factors which have little to do with military considerations' (ibid., p. 70), by which he means that the absence of such wars (the goal of general deterrence) may be due to non-military incentives and disincentives. But of course while military considerations inevitably loom larger in the outbreak of war they are still only one factor among many.