U.S. STRATEGIC NUCLEAR WEAPONS AND DETERRENCE

C. Johnston Conover

August 1977

P-5967
CONTENTS

Section

I. INTRODUCTION......................................................... 1

II. WHAT CAN STRATEGIC NUCLEAR WEAPONS DETER?...................... 3
    Type I Deterrence..................................................... 4
    Type II Deterrence.................................................... 7
    Type III Deterrence................................................... 9
    Other Types of Deterrence.......................................... 9
    Observations.......................................................... 10

III. HOW MUCH IS ENOUGH TO DETER?.................................... 13
    How Much Countervalue Capability Is Needed for Deterrence?.... 13
    How Much Counterforce Capability Is Needed for Deterrence?.... 19
    How Much Strategic Balance Is Necessary for Deterrence?....... 22
    How Necessary Is the Triad for Deterrence?...................... 26
    Observations.......................................................... 27

IV. HOW IS UNCERTAINTY RELATED TO DETERRENCE?........................ 30

V. HOW IS STABILITY RELATED TO DETERRENCE?............................ 36

VI. HOW IS FLEXIBILITY RELATED TO DETERRENCE?.......................... 41

VII. HOW CAN SALT CONTRIBUTE TO DETERRENCE?.......................... 49
    Preserving Mutual Assured Destruction............................ 49
    Limiting the Strategic Arms Competition.......................... 51
    Maintaining the Strategic Balance.................................. 53
    Observations.......................................................... 53

VIII. CONCLUSIONS......................................................... 55

BIBLIOGRAPHY.............................................................. 56
U.S. STRATEGIC NUCLEAR WEAPONS AND DETERRENCE

by

C. Johnston Conover*

The Rand Corporation
Santa Monica, California

I. INTRODUCTION

In the five years since the signing of the SALT I agreement, there has been a flood of new strategic literature which has sought to redefine, clarify and extend basic strategic concepts. In particular, much attention has been paid to how U.S. deterrence would work in a world in which strategic "parity" was a reality (the 1970s) or a world in which Soviet strategic "superiority" had come about (the 1980s). Deterrence is one of the few strategic concepts about which there is little ambiguity in definition. When the term "deterrence" is used in contemporary strategic literature, most analysts have in mind something similar to the following definition offered by George and Smoke [1974]: "In its most general form, 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." Likewise, most analysts could agree with a recent formulation by John Collins [1973] that deterrence encompasses "measures to prevent, rather than prosecute, wars, using psychological, as opposed to physical, means." Basically, deterrence has to do with persuading an enemy not to do what he would otherwise do. Under the broadest notion of deterrence, military forces alone are not the only avenue to its attainment. Kashi has recently delineated a number of approaches to

*Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The Rand Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The Rand Corporation as a courtesy to members of its staff.
deterrence which do not depend on military forces [Kashi, 1975]. However, this paper is devoted exclusively to the much narrower question of how strategic nuclear forces in particular contribute to U.S. deterrence against the Soviet Union.

The current disagreements about deterrence center not on the definition, but on basic questions such as these:

1. What can strategic nuclear weapons deter?
2. How much is enough to deter?
3. How is uncertainty related to deterrence?
4. How is stability related to deterrence?
5. How is flexibility related to deterrence?
6. How can SALT improve deterrence?

This paper is an attempt to synthesize current thinking about the role of strategic nuclear arms in deterrence in order to identify the cutting edges of disagreement and to clarify some issues in the debate. Although this paper analyzes only the most recent strategic literature which has contributed to the dialogue on deterrence, many of the basic notions about deterrence used in this paper and the debate are concepts originally introduced by analysts such as Kahn [1962; 1969]; Schelling [1966]; Wohlstetter [1959] and others. As with much of strategic literature, there is little being said currently about deterrence that has not been said before; there is probably little that was said before that has not been said in the current dialogue. Accordingly, for the convenience of the author in making this a tractable review, only the literature of the past five years has been included. Even here, this piece does not include every contribution to the debate; however most of the important works are included and it will be assumed that whatever is missing probably did not add an important new approach, concept or analytical tool for thinking about deterrence that is not already included in the literature reviewed here.
II. WHAT CAN STRATEGIC NUCLEAR WEAPONS DETER?

All strategic analysis of deterrence includes an implicit assumption about the types of enemy actions that can be specifically deterred by strategic nuclear weapons. As far as official U.S. doctrine is concerned, there has been a historical decline in the scope of enemy actions which nuclear weapons were intended to deter through declaratory policy. The "massive retaliation" notion of the 1950s was predicated on the assumption that the threat of massive retaliation could be used for Type I, Type II, and Type III deterrence.¹ Those who supported the assured destruction concept in the 1960s seemed to believe that Type I deterrence and Type II deterrence (in support of NATO) were still possible, but that the role of nuclear weapons in Type III deterrence was relatively small. The 1970s have seen increasing doubts about whether in a world of strategic parity or Soviet superiority, Type II deterrence is possible with just nuclear weapons. So although the usefulness of nuclear weapons in Type II and III deterrence has not been abandoned altogether, reliance on them is less now than in the 1950s.

Clearly, the extent to which deterrence operates depends in large part upon the relative size and balance of U.S. and Soviet forces: massive retaliation in the context of overwhelming U.S. strategic superiority has more plausibility as a Type III deterrent than it might in a world of strategic parity. The context of the current debate is one of rough parity in U.S. and Soviet forces in the 1970s and a possible margin of Soviet superiority in the 1980s (depending on the assumptions made about force capabilities and the particular measure of the balance used). So the real question being addressed in this section is: What can U.S. strategic nuclear forces deter in a world of strategic parity or slight Soviet superiority?

¹See Kahn [1969] for a definition of Type I, Type II and Type III deterrence.
TYPE I DETERRENCE

In considering the potential of strategic nuclear forces to achieve Type I deterrence, four basic scenarios are generally considered in the literature:

- An all-out Soviet attack on U.S. cities;
- An all-out Soviet attack on U.S. strategic forces, avoiding cities;
- A limited Soviet counterforce attack;
- A limited Soviet countervalue attack.

In addition, two levels of deterrence can be considered. The first level involves prevention of the attack; the second level involves "intrag- war deterrence" [Kahn, 1969]. Virtually all analysts agree that strategic nuclear weapons can deter an all-out Soviet attack on U.S. cities, since such an attack would leave the U.S. with overwhelming residual forces capable of inflicting massive damage on the Soviet Union. There is substantial disagreement about whether U.S. forces can deter a Soviet attack on U.S. strategic forces which leaves cities hostage to a follow-on strike. There are basically two lines of argument about why deterrence of this kind of attack might fail once Soviet ICBM and/or SLBM accuracies improve enough to be capable of destroying a substantial portion of U.S. ICBMs.

- Assured Destruction Is Infeasible -- some argue that Soviet civil defense and other damage limitation measures can keep population casualties and industrial damage quite low; hence, even with alert bombers and submarine forces for retaliation, the U.S. would be incapable of inflicting sufficient damage to deter the attack [Brennan, 1972b; Nitze, 1976b; Rummel, 1976].

- Assured Destruction Is Undesirable -- even if assured destruction levels were possible, the U.S. would be unwilling to retaliate since its cities are held hostage [Iklé, 1973; Nitze, 1976b; Rummel, 1976].
There are two major counterarguments used to explain why failure of deterrence in an all-out counterforce attack is not likely:

- **A Disarming Soviet First Strike Is Not Feasible** -- due to the synergistic effects of the triad [Downey, 1976] and fratricide, fusing problems and other technical constraints [Lodal, 1976; Steinbruner and Garwin, 1976], a large number of U.S. forces would survive.

  Even if only alert bombers and SLBMs survive, the U.S. would still have enough residual forces for assured destruction and limited options [Carter, 1974; Downey, 1976; Scoville, 1974], in particular if casualties beyond those due to prompt effects are considered [Panofsky, 1974].

  In the future, the likelihood of a first strike in a multilateral world is even less probable since the attacker would have to worry about retaliation by other nations besides the U.S. [Kaplan, 1975].

- **A Disarming Soviet First Strike Is Not Desirable** -- some analysts believe that the Soviets would have to be "absolutely certain" of eliminating U.S. retaliatory capability before launching a first strike [Cohen, 1973; Kahan, 1975]. A militarily significant counterforce attack would kill millions of Americans [Drell and Von Hippel, 1976; Scoville, 1974; Steinbruner, 1976]; therefore, even if the Soviets thought that the U.S. believed a countervalue retaliatory response would be suicidal, they would be uncertain about whether the U.S. might launch on warning or respond massively anyway; this risk is sufficient to maintain deterrence [Greenwood and Nacht, 1974; Lambeth, 1972; Lodal, 1976]. If the U.S. thought that reliance on Soviet uncertainty was insufficient as a deterrent, it could simply adopt a launch on warning policy whenever massive evacuation (which requires three days) was observed, and deterrence would work by denying the enemy his objective [Downey, 1976].
There is general agreement among analysts that limited counterforce and countervalue attacks can be in principle deterred using strategic nuclear forces. The major disagreement concerns whether such attacks require capabilities for similar options by the U.S. and whether a limited hard target counterforce capability is necessary to deter attacks on U.S. strategic weapons. This will be discussed at greater length in the section on deterrence and flexibility.

However, there is a major disagreement about the extent to which intrawar deterrence can be enhanced by nuclear forces in such contingencies. One school of thought argues that escalation is very hard to control for a variety of reasons:

- Misperceptions are certain to occur because of unintended collateral damage and other factors [Kahan, 1975; Steinbruner, 1976].
- Neither side would be able to contain its desire for revenge and would insist on inflicting the same or more damage on the other side as itself has already suffered [Downey, 1976].
- Even if political leaders were rational, there would be strong domestic pressures for escalation and it is questionable whether political leaders would be allowed to keep war limited [Cohen, 1973].
- Even if the U.S. could control its behavior in a limited nuclear war, Soviet doctrine is not geared to limited war, and even if it were, Soviet technical capabilities might not allow enough options to keep responses from escalating [Kahan, 1975].

Others, however, argue that U.S. restraint can induce restraint by the Soviets and that a capability to carry out limited nuclear options can be used to terminate conflicts arising from limited countervalue or limited counterforce strikes without escalation [Carter, 1974; Davis, 1976; Greenwood and Nacht, 1974], even in the event of Soviet surgical strikes against single targets [Lodal, 1976; Schlesinger, 1975]. However, within this group, there is disagreement about whether such intrawar
deterrence requires an improved U.S. hard target capability. Some argue that current capabilities provide sufficient counterforce options [Carter, 1974; Greenwood and Nacht, 1974]; however, others argue that improvements in accuracy and yield are warranted to extend the range and efficiency of U.S. counterforce options [Lodal, 1976; Schlesinger, 1974; Van Cleave and Barnett, 1974]; only then will deterrence against the full range of possible Soviet limited strikes be assured.

**TYPE II DETERRENCE**

Most analysts believe that with a margin of U.S. superiority, the use of strategic nuclear weapons as an umbrella is a credible deterrent (although some argue that even with superiority, the U.S. did not deter Soviet moves in Berlin and Cuba [Bailey, 1973]); some argue that strategic superiority is necessary to offset NATO conventional force inferiority [Rosecrance, 1975]. There is disagreement about how credible the umbrella remains in a world of strategic parity. When parity exists, some argue that Type II deterrence credibility is low [Lambeth, 1972; Nitze, 1976b; Rosecrance, 1975], while others argue that the role of strategic forces in deterrence merely diminishes relative to the role of conventional forces, but the umbrella remains intact [Schlesinger, 1974]. Another view is that superiority is not necessary for Type II deterrence so long as U.S. resolve and commitment to nations under the umbrella is clear [Lambeth, 1972].

In a world of parity, what can be done to maintain Type II deterrence? Under certain circumstances, if credibility is low, it can be offset by increasing the deterrence threat but beyond a certain point, adding to the threat adds little to deterrence [Rosecrance, 1975]. At current levels of forces, most would agree that adding to the threat will not bolster the credibility of the umbrella. Four possibilities have been suggested for improving Type II deterrence credibility in a world of parity:

- **Limited Nuclear Options — LNOs can be used to demonstrate resolve and to increase the risk to the Soviets that escalation may occur [Davis, 1976; Schlesinger, 1974].** William
Beecher argues that a limited U.S. hard target kill capability would be useful in deterring Soviet use of MRBMs in Europe [Martin, 1974]. In addition, it is argued that an assured destruction threat would not be credible in many limited NATO scenarios (such as an attack on NATO tactical nuclear weapon storage sites) [Iklé, 1973; Lambeth, 1976]. LNOs are a means of overcoming this credibility gap.

- Rely on Uncertainty -- even without LNOs, some believe that Soviet uncertainty about escalation is sufficient to maintain deterrence [Kahan, 1975], and Europeans in particular seem to rely heavily on uncertainty as a key factor that maintains deterrence [Davis, 1976; Martin, 1977] (however, some Europeans do favor LNOs as an additional means of enhancing deterrence [Martin, 1977]).

- Improve Conventional Forces -- some analysts have noted that if too much reliance is placed on the nuclear umbrella and conventional forces are reduced, this weakens deterrence by making the NATO commitment look weak [Rosecrance, 1972]; Howard argues that once conventional forces are actually engaged in combat, the credibility of nuclear use becomes much stronger [1973]; therefore, even if nuclear forces cannot deter the initiation of an attack, they are still useful and credible for intrawar deterrence.

- Improve U.S. Strategic Defense -- some argue that Type II deterrence credibility would be considerably stronger if it were perceived that the U.S. could survive a nuclear war and that damage limitation improvements would therefore be desirable [Bailey, 1972].

In short, Type II deterrence is feasible under conditions of parity (even for limited attacks) but it is less important, and many analysts believe it may require the addition of capabilities such as LNOs, improved conventional forces or improved U.S. damage limitation capabilities in order to remain credible.
TYPE III DETERRENCE

William Kaufman has stated that deterrence credibility depends in part on cost: there should be some relationship between the value of the objective sought and the costs involved in seeking it [Rosecrance, 1975]. Given the risks involved in crossing the nuclear threshold, this suggests that the use of strategic weapons for Type III deterrence will be inherently less credible than for Type II deterrence. On the other hand, Kahn's original notion of Type III deterrence focused on limited "tit-for-tat" responses [Kahn, 1969], and some argue that such responses would be inherently more credible than the threat of an all-out spasm response. The issue is whether using even limited options for provocations other than direct attacks on the U.S. or its allies would be credible in a world of parity or Soviet superiority, insofar as in such worlds the U.S. might no longer have escalation dominance. Very little has been said in the current debate about this issue. Davis suggests that proponents of limited nuclear options believe that LNOs could be used to deter "the use of all kinds of military action . . . anywhere in the world" [Davis, 1976], but it is not clear whether this refers only to actual military attacks (in which case Type II deterrence is relevant) or to more limited uses of military forces as well. Schlesinger indicated in 1974 that in addition to deterring conventional and nuclear attacks on allies, U.S. strategic forces were viewed as "helping to inhibit the coercion of our allies" by nuclear powers [1974], but no explanation is provided as to what contingencies are covered by "coercion." It suggests, however, that the use of strategic nuclear forces for Type III deterrence was contemplated and thought to be feasible.

OTHER TYPES OF DETERRENCE

A number of analysts have noted that nuclear weapons are useless in deterring accidents, unauthorized launches, irrational leaders or escalation of crises by third parties [Gray, 1973; Ikle, 1973; Panofsky, 1974]. On the other hand, LNO advocates argue that intrawar deterrence is feasible in such instances to control escalation after the inadvertent initiation of nuclear war [Schlesinger, 1974]. However, others argue that flexible options increase the chances for an accidental launch [Ball, 1974] and
that the best way to reduce the likelihood of accidental or unauthorized use is to reduce the number of nuclear weapons available [Panofsky, 1974]. In addition, some believe that in the future the prospects for intrawar deterrence after undeterrable launches will decrease, since in a multilateral world, the incentives to strike back will be lower for the aggrieved party for fear that retaliation would provoke a response by other nuclear nations who would regard retaliation as irresponsible and who would launch a counter-response to punish the offending nation [Kaplan, 1975].

One recent novel notion of what strategic nuclear forces can deter concerns enemy force deployments. Steinbruner argues that the U.S. has implicitly been seeking to deter certain Soviet force deployments by "the threat of retaliation, even preemption—in terms of weapons development and deployment; this has become "a major if not dominant element of the United States bargaining strategy." [Steinbruner, 1976]. Steinbruner questions whether such secondary deterrent strategy can work, but if what he says is true, the U.S. defense and arms control bureaucracy has contributed a unique version of deterrence for consideration in the debate.

OBSERVATIONS

As others have observed [Greenwood, and Nacht, 1974], there is too much preoccupation in current U.S. strategic thinking with all-out strategic attacks and too little attention paid to more likely scenarios involving strategic nuclear weapons such as NATO contingencies or limited nuclear attacks in crisis. Even when MARV accuracies are achieved, the prospect of a Soviet first strike or blackmail threat seems quite unlikely. Two questions need to be asked.

First, who is at a bargaining disadvantage when the Soviets evacuate their cities and demand concessions from the U.S.? The U.S. either could announce a launch on warning posture (as Downey suggests) or it could simply claim it will ride out the attack and respond with SLBMs and bombers as soon as the evacuation is over. The analysts who posit this blackmail scenario mistakenly assume that speedy retaliation is essential for deterrence and they ignore the fact that evacuation cannot last forever.
The second question is: If the Soviets did launch an all-out counterforce attack, and the U.S. responded with a bomber strike to mop up all Soviet ICBM silos to preclude reloading, who is at a bargaining disadvantage? As the bombers performed this retaliatory strike, presumably the Soviets also would be deterred from striking U.S. cities since U.S. SLBMs would be held in reserve to keep them hostage (if this logic applies to the U.S. side, it must, for consistency, apply to the Soviet side as well). Given that the Soviets have so much more of their strategic forces in land-based silos, the military balance after such an exchange would overwhelmingly favor the U.S. What would be the incentive for such an attack?

As will be explained in detail later, if the Soviets are in a "lose now, lose later" situation, the likelihood of a desperate all-out attack becomes greater. But in such a situation, it does not seem likely that even a massive improvement in U.S. forces would provide any more of a deterrent than current and projected capabilities. In the "lose now, lose later" situation, the leaders of a nation believe that without attacking, the nation will lose in the long run almost with certainty. However, even if the outcome of a war looks dismal, there is enough uncertainty about the outcome of war that an attack is warranted: in such a situation, the leaders prefer the risk of losing in the short run to the certainty of losing in the long run. It is difficult to imagine a strategic nuclear force configuration or declaratory posture that would deter an attack in such circumstances. For example, even if the Soviets knew for certain that the U.S. would launch everything on warning, the uncertainty ranges for nuclear war outcomes are so vast that the risk would still be worth taking relative to the certainty of long-run defeat. However, the limitations of this approach will be discussed in Section III.

In limited attack scenarios, the concern about self-deterrence seems valid if U.S. responses were limited to fairly large attacks. However, the validity of that concern does not translate into an argument for improved hard target kill capabilities. There do not appear to be any valid military grounds for why a U.S. response to a Soviet attack on an ICBM
silo would have to be perfectly symmetrical in order to be a response in kind. Would not a strike against another military target be an equally valid response in kind? Even if a symmetrical response in kind were desired, why does it matter that it would take two or three Minuteman missiles to take out a Soviet silo rather than only one? Schlesinger has not satisfactorily answered the question of why improved efficiency of current U.S. hard target kill capabilities is necessary for deterrence in such scenarios.

Finally, as for Type II deterrence as it relates to NATO in particular, reliance on a combination of improved conventional forces and limited nuclear options appears to be the best means of upgrading the credibility of the umbrella in a world of nuclear parity. Limited nuclear options help remove the doubt about whether the U.S. would commit suicide on behalf of West Germany. Conventional force improvements help bolster perceptions of NATO resolve and that ultimately improves the credibility of the nuclear deterrent. More importantly, both of these offer intrawar deterrence advantages in the event that deterrence fails. LNOs are a credible and militarily useful response to conventional attacks in Europe (unlike assured destruction or its relatively large suboptions) and their credibility is enhanced even further after conventional forces are engaged. There may be other policy reasons that the U.S. would not want to rely on Type II deterrence anymore, but if it does, a combination of these two options would allow Type II deterrence to continue to be credible even under Soviet superiority (within a limited range).

In short, strategic nuclear weapons have the potential to continue providing primary and intrawar deterrence of a wide range of possible Soviet actions, ranging from all-out attacks on the U.S. to even limited attacks in other areas of the world -- even under conditions of parity or the margin of Soviet superiority projected for the 1980s.
III. HOW MUCH IS ENOUGH TO DETER?

In using deterrence notions as criteria for sizing U.S. strategic forces, four major issues are discussed.

- How much countervalue capability is needed for deterrence?
- How much counterforce capability is needed for deterrence?
- How much of a balance in U.S. and Soviet strategic forces is necessary for deterrence?
- How necessary is the triad for deterrence?

Most analysts agree that some countervalue capability is needed for deterrence, but disagree about how much. There is considerable disagreement about whether any counterforce capability is needed, given an assured destruction level of countervalue capability. There is also disagreement about whether the strategic balance makes much difference in the capability or willingness to carry out a deterrent threat. Finally, there is a controversy over how useful the triad is in maintaining deterrence.

HOW MUCH COUNTERVALUE CAPABILITY IS NEEDED FOR DETERRENCE?

Absolute Criteria

The first attempt to precisely define the levels of nuclear weapons necessary to implement a continuing Type I deterrent threat came under McNamara. In his 1965 posture statement, McNamara set the necessary requirements for assured destruction as the capability to destroy between one-quarter to one-third of Soviet population and two-thirds of Soviet industry [Russett, 1974]. By 1967, he had changed this criterion to one-fifth to one-fourth of Soviet population and one-half to two-thirds of Soviet industry [McNamara, 1967]. Finally, in his final posture statement made in 1968, McNamara further lowered the assured destruction requirements to one-fifth to one-quarter of population and one-half of Soviet industry [Holst, 1969]. A surviving force of 400 one-megaton equivalents was deemed sufficient to meet this deterrence requirement [U.S. Congressional Budget Office, 1977]. Since McNamara resigned, no
subsequent Secretary of Defense has tied U.S. deterrence requirements
to a specific absolute minimal level of destructive capability:

- Clark Clifford said that the assured destruction capability
  of U.S. strategic forces was measured in terms of "their
  ability, even after absorbing a well-coordinated surprise
  strike, to inflict unacceptable damage on the attacker"
  [1969];
- Melvin Laird indicated that deterrence required "an adequate
  second strike capability" [1971];
- Donald Rumsfeld recently stated that "assured retaliation"
  required a capability of inflicting "irreparable damage" on
  Soviet industry, population and military targets [1977].

Among strategic analysts, a variety of absolute deterrence criteria
has been proposed, but most of these are similar to McNamara's criteria:

- Kahan argues that the guaranteed destruction of 20-25 percent
  of population and 70 percent of the Soviet industrial base
  is more than sufficient as a deterrent [1975];
- Downey believes that the assured destruction of one-third of
  population and two-thirds of industry is sufficient for deter-
  rence [1976];
- Quanbeck and Wood claim that destruction of one-third of popula-
  tion and three-quarters of industry is enough [1976].

How do these criteria translate into force requirements? Downey
suggests that delivery of between 200 to 400 one-megaton weapons would
be sufficient to meet his deterrence requirements and that possession of
deliverable amounts below this range is insufficient and amounts above
this range are hypersufficient [1976]. Quanbeck and Wood say that 400
one-megaton equivalents can achieve their deterrence criteria [1976].
Others have speculated that while 400 megaton equivalents is necessary
currently for deterrence, U.S.-Soviet relations could improve in the
future to a point where only 50 Minuteman ICBMs would be a sufficient
deterrent [Ruina, 1973].
The attitude of most of the above analysts is summarized well in the following quote:

If the Soviet Union were not deterred at the prospect of the arrival of 400 megatons of destruction, it scarcely seems likely that she would be fazed by an additional two to three hundred bombs. [Rosecrance, 1975]

The notion that deterrence requires only a certain minimal level of assured retaliatory capability against cities is known as "finite deterrence" [Kahn, 1969; U.S. Congressional Budget Office, 1977]. Those who disagree with advocates of finite deterrence raise four arguments:

- Even if assured destruction criteria were enough for deterrence, a force level of 200 to 400 megaton equivalents is too small to achieve even the lower bound of McNamara's criteria (one-fifth of population; one-half of industry). Nitze claims that a much larger second strike residual force (at least 3000 deliverable megatons) is needed to achieve assured destruction levels of damage [1976b].

- Assured destruction capability is not enough: counterforce capabilities and a balance of forces are also critical in maintaining deterrence [Nitze, 1976a; Rumsfeld, 1977; Schlesinger, 1974].

- Assured destruction is dangerous: in order to maintain a credible deterrent, a firm policy of assured destruction may actually lead to its being executed after an attack and that would be completely irrational [Russett, 1972].

- Deterrence depends on context: a nation whose long-term prospects are bleak may be willing to risk a larger amount of damage than if prospects are brighter [Brown, 1977; Snyder, 1976]; hence, the level of retaliatory damage needed for deterrence is a "variable, not a constant" [Russett, 1972].

There is considerable ambiguity in the literature regarding the extent to which absolute criteria for Type I deterrence also allows an umbrella effect for Type II and Type III deterrence. Ian Smart claims that Europeans generally believe that a capability to deliver 100 to
200 weapons above and beyond those needed for assured destruction is sufficient. It has elsewhere been argued that any umbrella effect must rely on forces in surplus of the minimal level of force needed to threaten Soviet hostages (since this guarantees immunity for the U.S. population) [Martin, 1974].

This review would not be complete without mention of those who believe that no counter-city destruction is necessary for deterrence. These analysts believe that targeting cities is immoral since assured destruction cannot guarantee that nuclear war will not occur (due to undeterrable events such as accidents) and since other options are available to maintain deterrence besides counter-city threats [Brennan, 1972b; Gray, 1973; Iklé, 1973; Van Cleave and Barnett, 1974]. These critics of assured destruction have proposed two alternatives to mutual assured destruction:

- No Cities Targeting -- some prefer to target strategic military forces and military-related industry [Iklé, 1973], while others would prefer to also target police and conventional forces in order to render the Soviets helpless against internal uprisings and Chinese and Eastern European troops [Russett, 1972, 1974]; Iklé argues that if counter-city targeting is thought to be desirable, the emphasis on speedy retaliation should be eliminated since it is not necessary for deterrence; this would at least lessen the probability of accidental launch [1973].

- Damage Limitation -- others argue that mutual assured destruction guarantees that nuclear war would be a disaster for all -- even if caused by undeterrable events [Bailey, 1972; Brennan, 1972b]; they believe that more emphasis on population defense, including ABM and civil defense, could be implemented without degrading deterrence [Bailey, 1972; Brennan, 1972b; Coffey, 1973; Iklé, 1973].

These views are countered by those who regard mutual assured destruction as the best that can be achieved given that defense against massive nuclear attack is impossible [Panofsky, 1974; York, 1973]. Some
argue that moving to "no cities" targeting would make nuclear war more attractive and would substantially increase the probability of war [Kahan, 1975]. They regard ABM and good civil defense as destabilizing insofar as these would degrade Soviet retaliatory capabilities and might heighten fears of a Soviet first strike [Kahan, 1975; Perle, 1973; York, 1973]. In addition, they claim that large scale defensive systems would make the balance very unstable since small violations in SALT or other agreements would be much more significant [Kahan, 1975]. Some also have noted that large-scale defense systems would preclude limited options since attacks would have to be relatively large for even a few weapons to land on target [Kahan, 1975; Scoville, 1974]. However, some proponents of mutual assured destruction do acknowledge that certain features of the current U.S. posture could be eliminated without impairing deterrence -- such as the emphasis on rapid retaliation [Cohen, 1973; Panofsky, 1974].

Relative Criteria

In recent years, more importance has been attached to relative criteria in assessing the forces necessary for deterrence. Both McNamara and Clifford relied on assured destruction as the primary criterion for determining whether deterrence credibility could be maintained (in 1968, however, Harold Brown, then Secretary of the Air Force, argued that a less central but still important criterion for deterrence should be that "ratios of surviving population and industry must not be badly adverse to the United States" [Newhouse, 1973]).

Under Nixon, the concept of assured destruction was retained but a broader notion of "sufficiency" was advanced which included a criterion quite similar to Brown's: the U.S. sought to maintain an ability to prevent "the Soviet Union from gaining the ability to cause considerably greater urban-industrial destruction that the United States could inflict on the Soviets in a nuclear war" [Laird, 1971]. The sufficiency criteria also included a requirement that the U.S. provide "no incentive for the Soviet Union to strike the United States first in a crisis" [Laird, 1971]. This relates to a concept advanced by Daniel Ellsberg in the early 1960s and widely accepted in the current generation of analysts that an important element of stability is that the difference in first-strike and second-strike outcomes be minimized [Kahan, 1975; Perle, 1973; Rosecrance, 1975].
These criteria suggested that maintenance of crisis stability was an important part of deterrence and that in addition to assured destruction, a sufficient deterrent posture would have to ensure that the Soviets did not develop incentives for a first strike. Such incentives could arise if the Soviets perceived:

- A significant difference between expected Soviet counter-value losses in a first strike compared to a second strike;
- A significant difference between expected Soviet counter-value losses and U.S. countervalue losses.

Such incentives could be curbed only by maintaining the appropriate balance of forces and capabilities, but the Nixon administration never proposed any force size requirements to implement these relative criteria.

Under the impetus of Schlesinger, these relative criteria were expanded in NSDM-242 in April 1974. This Memorandum included an "anti-recovery" concept that Rumsfeld later translated into a deterrence criterion as follows: the U.S. "assured retaliation" forces must include a capability to "retard significantly the ability of the USSR to recover from a nuclear exchange and regain the status of a 20th century military and industrial power more rapidly than the U.S." [Rumsfeld, 1977]. According to the Congressional Budget Office, the force requirements to implement the relative recovery criterion would be "much more demanding" than those of a minimum deterrence policy [1977]; indeed, Rumsfeld's Defense Department Report suggests 8500 warheads would be needed to sufficiently cover the targets necessary for implementing anti-recovery [1977]. However, it has been recently reported that Harold Brown is moving away from the anti-recovery concept because it is too ambiguous and could be used to justify enormous increases in strategic forces. In addition to Brown's dissatisfaction with the anti-recovery notion, the entire idea of relative countervalue criteria is challenged by some analysts who question whether relative differences even matter in an all-out attack [Kahan, 1975; Russett, 1972]. This critique is most often leveled by finite deterrence advocates who argue:

beyond the ability of either side to inflict 75 million fatalities and between 50 and 75 percent industrial damage -- levels that would finish either country as a viable society --
relative differences in the ability to inflict urban or industrial damage seem insignificant [Carter, 1974].

**HOW MUCH COUNTERFORCE CAPABILITY IS NEEDED FOR DETERRENCE?**

**Absolute Criteria**

Unlike countervalue criteria, there has been little discussion of adopting a minimal level of counterforce capabilities as a criterion for deterrence sufficiency. One possible exception to this might be Russett, who has argued for a "countercombatant" alternative to assured destruction [1972]. In addition to targeting strategic forces and military-related industry (as others have suggested), Russett would target Soviet police and Soviet military troops on the borders of Eastern Europe and the People's Republic of China; this strategy would then make the Soviets vulnerable to threats against which its remaining strategic forces would be relatively useless -- including internal uprisings and conventional threats from Eastern Europe and China [1974]. Russett argues that although Soviet strategic forces would be targeted, there would be no need to destroy any particular level or proportion of Soviet weapons [1972]. By implication, then, Russett's strategy would necessitate a fixed level of U.S. counterforce capability; he has not specified the precise level necessary, but has indicated that since the countercombatant strategy involves more targets than a pure assured destruction capability, more forces would be required than for assured destruction requirements [1972].

Schlesinger's argument for limited nuclear options (what Collins would call "counterforce as insurance" [1973]), implies that some minimal level of counterforce capability is necessary to carry out limited attacks. However, Schlesinger never specified the minimal amount of counterforce capability necessary. In addition, because of his emphasis on maintaining similar limited options capabilities, Schlesinger's LNO requirement for deterrence probably more appropriately belongs under relative criteria.
Relative Criteria

Nitze and Schlesinger are the foremost proponents of maintaining a balance between U.S. and Soviet counterforce capabilities. Nitze would like the balance of forces to be such that by striking first, the Soviets could not improve their net throw-weight ratio over the U.S. either in terms of the ratio of Soviet to U.S. throw-weight or in terms of the net difference between Soviet and U.S. throw-weight (measured in pounds). Nitze estimates that for the mid- to late-1980s, this would require a capability to destroy 1500 to 2000 hard targets with high probability (although this requirement would change with Soviet forces) [1976b].

A more general case for maintaining a balance in counterforce capabilities was offered by Schlesinger, who argued that the best deterrent of any Soviet attack on U.S. ICBMs would be a capability to respond in kind [Davis, 1976]. Further, Schlesinger indicated that "we do not intend that the Soviet Union should have a wider range of options than we do" and that "we do not propose to let an opponent threaten a major component of our forces without our being able to pose a comparable threat" [1974]. These suggest that Schlesinger was arguing for a relative balance of counterforce capabilities -- i.e., that the U.S. would match the scope of Soviet counterforce deployments and not just match types of capabilities.

In contrast, there are others who believe that the U.S. should not permit the Soviets to obtain a significant military capability that the U.S. does not have (including limited nuclear options), but who also argue that U.S. residual forces even after an overwhelming first strike would still be sufficient for deterrence [Lodal, 1976]. Hence, there is no particular need to maintain a balance in counterforce capabilities.

Another rationale for maintaining some sort of balance in counterforce capabilities would be to maintain essential equivalence. Van Cleave and Barnett claim that maintenance of essential equivalence would require a larger hard target kill capability than would Schlesinger's requirement for more flexible targeting -- but they never specify the precise level desired [1974].
The general argument against preserving relative counterforce capabilities is that development of U.S. counterforce capabilities threatens stability (an immediate threat to deterrence) and in the long run may harm SALT prospects (and thereby presumably would have a long-run impact on deterrence). Some argue that there is no way to distinguish forces intended for a first-strike counterforce from those for retaliatory counterforce. Hence, even the attainment of very limited U.S. counterforce options would be interpreted by the Soviets as an attempt to get a first-strike and this would be destabilizing [Kahan, 1975; Panofsky, 1974]. This suggests that any benefits from a balance of counterforce capabilities would be more than offset by the costs of instability and that deterrence would be weakened, not strengthened, by U.S. counterforce capabilities. This issue will be dealt with in more detail in the section on deterrence and stability.

A more extreme critique is offered by Downey, who argues that there simply are no benefits at all to be obtained from U.S. counterforce capabilities. The argument is that the essential purpose of strategic forces is to deter Soviet aggression by the threat of countervalue retaliation (since adequate defense to deny Soviet objectives in a first-strike is impossible to achieve). Therefore, counterforce capabilities are significant only to the extent that they could degrade enemy capabilities to destroy one's own countervalue targets. U.S. counterforce capabilities would be useful if the U.S. intended to initiate a nuclear war, but as a deterrent (i.e., a threatened response), "to respond to an opponent's counterforce capability by creating one's own counterforce capability is an exercise in irrelevance since it does nothing to reduce the impact of his counterforce upon one's own countervalue" [Downey, 1976]. Downey concludes with certain other analysts that, even if the Soviets have superiority in counterforce capabilities, this would not be militarily useful so long as the U.S. retained a capability for retaliation on countervalue targets [Downey, 1976; Lodal, 1976; Steinbruner, 1976].
HOW MUCH STRATEGIC BALANCE IS NECESSARY FOR DETERRENCE?

There are basically four positions which can be taken on the question of how the strategic balance (however measured) affects deterrence:

- Any margin of Soviet superiority in the balance of strategic forces degrades deterrence;
- A small margin of Soviet superiority in the balance of strategic forces will not affect deterrence;
- Only a large margin of Soviet superiority in the balance of strategic forces will affect deterrence.
- The strategic balance is irrelevant for deterrence.

There are three important arguments made by those who believe that even small strategic imbalances may diminish the deterrent value of U.S. forces:

- Chances for U.S. Capitulation Increase -- some argue that in crisis situations, the degree of Soviet superiority is irrelevant; leaders on both sides will form expectations and shape behavior in accordance with which side they believe will come out ahead in the end [Brennan, 1972a]. Since the maintenance of peace depends on the accommodation of expectations, there would be strong pressures on the weaker side to concede something [Kaplan, 1975]. In addition, "under conditions of nuclear parity, the power which can force upon its adversary the decision to initiate the use of nuclear weapons enjoys an enormous strategic advantage" since political leaders may constrain military personnel from taking action in such situations [Howard, 1973]; given the conventional imbalance, without offsetting margin of U.S. strategic superiority, the Soviets might have this kind of strategic advantage in a NATO contingency; on grounds such as these, many have argued that at least some degree of U.S. strategic superiority is necessary to maintain the U.S. umbrella [Rosecrance, 1975]; or
more generally to deter Soviet power projection anywhere on its periphery [Nitze, 1976b].

- Chances for Soviet Miscalculation Increase -- advocates of essential equivalence argue that asymmetries in the various measures of the balance are tolerable if "they do not all favor one party ... we are not prepared to accept a situation in which all the visible asymmetries point in one direction" [Schlesinger, 1974]; one concern is that with superiority, the Soviets might be misled into thinking they could exploit their advantages for diplomatic purposes and initiate a crisis [Schlesinger, 1974]; at minimum, Soviet aggressiveness and tenacity in crises would increase if the balance were generally perceived to be in their favor [Brennan, 1972a; Van Cleave, 1973].

- Chances for Allied Accommodation Increase -- some worry about "general Soviet hegemony," and accommodation [Nitze, 1974] or "Finlandizing" tendencies among U.S. allies [Brennan, 1972b] if the Soviets attain superiority; although critics believe that the degree of balance is irrelevant to NATO allies, others claim that European perceptions of the balance and estimates of relative strategic superiority and inferiority have been quite important in past European calculations about limited nuclear options [Martin, 1977].

Those who believe that some margin of Soviet strategic superiority would not seriously degrade deterrence use two major lines of reasoning:

- Slight U.S. Inferiority Can Be Overcome -- some claim that a strong U.S. political posture can diminish the value of a slight symbolic Soviet strategic superiority [Kaplan, 1975]. Similarly, since the U.S. largely helps to create world perceptions, the U.S. can minimize political problems created by disparities in static measures of the strategic balance simply by refusing to attach importance to them [Lodal, 1976]. Others claim that a complex web of
relationships affects perceptions, so that the U.S. can comfortably accommodate a slight margin of inferiority and even maintain a credible umbrella so long as "major asymmetries favoring the Soviets are avoided [Greenwood and Nacht, 1974].

Significant U.S. Inferiority Is Dangerous -- Samuel Huntington claims that if the Soviets did achieve "clear superiority," the American people simply would not tolerate it and would inevitably seek to regain equality; this renewed arms race would make the risk of war great and is one reason why major imbalances should be avoided [Ntze, 1974].

A third school of thought believes that only a large margin of Soviet superiority would have an impact on deterrence [Lambeth, 1972; Quanbeck and Bleckman, 1973; Steinbruner, 1976]. Two important arguments underlie this belief:

- Any strategic advantage short of a credible first strike offers no direct coercive advantage; a credible first-strike is technologically infeasible, therefore, a large Soviet advantage is of no consequence [Lambeth, 1972];
- In crisis, the technical details of the balance are irrelevant to decisionmakers; they can attend only to the massive destructiveness of even a few nuclear weapons. Therefore so long as the U.S. maintains a survivable deterrent, relative balance does not matter [Steinbruner, 1976].

Lambeth's argument is that asymmetries in interests, commitments and levels of resolve are far more important in determining crisis outcomes than asymmetries in strategic forces. In addition, when both sides hold a mutual assured destruction capability, any advantage in strategic weapons cannot be used for coercive advantage (unless one side holds a credible first-strike capability) since any threat to use nuclear weapons will be incredible [Lambeth, 1972].
Steinbruner's argument is underpinned by his cybernetic theory of decisionmaking. Under this theory, even with a strong advantage, the decisionmaker would make very conservative assumptions in attack planning and would not be concerned with the degree of a retaliatory threat (within a broad range); he would simply expect retaliation or not expect it and that is all that matters for deterrence. Consequently, Steinbruner argues that the U.S. should tolerate "substantial disparities" in static force measures of the balance so long as an undeniable second strike retaliatory capability is maintained. In response to the argument that asymmetries in the balance generate expectations that ultimately translate into disadvantageous political effects for the weaker side, Steinbruner comments that "the hypothetical political effects of these disparities range from highly speculative to highly implausible" [Steinbruner, 1976].

For the future, some argue that in a multilateral world, the importance of the U.S.-Soviet strategic balance will diminish considerably since the existence of many actors will make a disarming first strike a less credible threat [Kaplan, 1975; Pfaltzgraff, 1973]. The use of strategic superiority to extract political gains is thought to be less likely since an aggressor would have to be concerned with the reactions of other nuclear nations [Pfaltzgraff, 1973]; it is believed that other nations would band together to punish any irresponsible actions of an aggressor on grounds that they may become the next victims if the aggressor's irresponsibility is not curbed immediately. Even if such nations in fact would not band together, some claim there would be enough uncertainty about this prospect to provide a deterrent effect [Pfaltzgraff, 1973].

Finally, there are those who argue that the strategic balance is largely irrelevant to deterrence. Some of these question how large margins of strategic superiority could realistically be translated into political advantage and this is reflected in Kissinger's well-known lament: "What in the name of God is strategic superiority? What is the significance of it politically, militarily, operationally at these levels of numbers? What do you do with it?" [Nietze, 1974]. Others argue that the
decision to go to war is dominated by factors other than the state of the balance — the most important factor being the leadership's perceptions of the long term political, economic and social prospects for the nation [Brown, 1977]. If the long-term prospects look bleak, a nation may choose to go to war regardless of an unfavorable balance simply because war is perceived as the least miserable option [Brown, 1977; Perle, 1973]. In such cases, the leaders prefer the uncertainty of losing a war to the certainty of losing in the long run if it does not attack. Conversely, a nation with bright long-term prospects might not initiate war even if the balance is overwhelmingly favorable.

**HOW NECESSARY IS THE TRIAD FOR DETERRENCE?**

The U.S. triad has been justified on the basis that it guarantees the survival of assured destruction forces for deterrence, that it offers flexibility for strategic planning and that it offers a hedge against technological breakthroughs which might make a particular arm of the triad vulnerable [Ruina, 1973]. Because of the triad, with current technology and strategic forces, it is not possible for either the Soviet Union or U.S. to launch a disarming first strike which simultaneously eliminates all ICBMs, SLBMs and bombers [Downey, 1976; Lambeth, 1972; Quanbeck and Blechman, 1973], and this assures that deterrent forces would remain for retaliation. However, analysts do not agree on exactly how the triad does or should relate to assured destruction requirements or to other criteria for assessing the adequacy of deterrence.

For example, Perle argues that it is U.S. policy to maintain a deterrent capability in each of the components of the triad to ensure that loss of two would not leave retaliatory capability impaired [Perle, 1973]. Quanbeck and Blechman appear to imply the same when they state "an independent retaliatory capability is maintained in the three separate offensive systems . . . that compose the Triad" [Quanbeck and Blechman, 1973]. However, Schlesinger argues that the U.S. purpose in maintaining a triad "is not to provide an independent assured destruction capability in each element of the strategic forces. Rather, it is to achieve a sufficient degree of diversification in our forces to hedge against both foreseeable and unforeseeable risks" [Schlesinger, 1974].
Regardless of doctrine, each arm of the current triad of forces does carry independent countervalue capabilities in excess of the requirements for assured destruction laid out by McNamara [Downey, 1976]. Some argue that a triad is the minimum necessary as a hedge against technological breakthroughs. Thus, Downey has recently stated "there is, of course, nothing divinely ordained about a three-armed deterrent: four or five independent arms would be better (if such a thing were technologically possible, which to the best of my knowledge it is not); fewer arms would be worse" [1976].

However, the necessity of maintaining a triad in the long term has been questioned by a number of analysts. Thus, in opposing production of the B-1 bomber, the Center for Defense Information (and others) have questioned whether a manned bomber is needed past the mid-1980s [Center for Defense Information, 1973; 1975]. Similarly, Quanbeck and Blechman argue that land-based missiles should be eliminated over the long term to enhance stability [Quanbeck and Blechman, 1973]. Others argue that moving to a dyad would weaken the credibility of deterrence since it would diminish diversity in strategic forces; in addition, the Soviets might try to exploit a monopoly in land-based missiles if the U.S. unilaterally moved to a bomber-SLBM dyad. In this view, a negotiated phase-out of ICBMs on both sides would be preferable, although if such negotiations failed, then a unilateral phase-out of ICBMs might be warranted [Kahan, 1975]. The sea-based forces are generally regarded as the most survivable component of the triad and some advocate increasing reliance on sea-based forces for deterrence rather than on other strategic arms [Scoville, 1972b; Quanbeck and Wood, 1976]. Thus, there is controversy about whether more or less than a triad is needed, but no serious discussion in recent years has focused on relying solely on a strategic nuclear monad to provide deterrence; the consensus even among finite deterrence advocates appears to be that a monad would be undesirable [Downey, 1976; Kahan, 1975].

OBSERVATIONS

In discussing the requirements for deterrence, it seems that many analysts are talking past each other because of differing conceptions of what strategic weapons are intended to deter.
The use of finite deterrence criteria focuses too heavily on the feasibility of inflicting pain in response to a Soviet first-strike and not enough on the credibility of such a response. Maintaining forces to meet such criteria is useful only for deterring an all-out attack on cities -- but that seems to be the least likely scenario by which nuclear war might come about. Advocates of finite deterrence implicitly (and sometimes explicitly) assume that even though assured destruction might not be credible for much more limited nuclear attacks, NATO contingencies or other Soviet actions, reliance on Soviet uncertainty about a U.S. response is sufficient. In a question as important as nuclear war, is it really desirable that so much reliance be placed on uncertainty? I believe not, and therefore think that additional forces and capabilities are needed to deter those actions against which the assured destruction threat is relatively impotent.

Limited nuclear options make deterrence threats credible and offer intrawar deterrence that is not possible with pure assured destruction forces. Such flexibility does not require improved hard target kill capabilities, however. The primary usefulness of limited nuclear options lies in demonstrating resolve -- not in warfighting per se. Any LNO capability designed for significant damage limitation would require a counterforce capability so large as to be destabilizing: the mutual possession of counterforce capabilities large enough to threaten very large portions of each side's land-based missiles would lead to crisis instability and preemption incentives that are undesirable. There are sufficient soft military targets available for such demonstrations of resolve without the danger of a lot of collateral damage. Furthermore, even if response in kind were thought to require an ICBM kill capability, that can currently be achieved with two or three Minuteman IIIs. It seems improbable that the use of three ICBMs for such limited attacks would be perceived as more belligerent or escalatory than the use of one ICBM; hence, it is not at all clear why more "efficient" hard target kill capabilities would be needed for this kind of deterrence.

The controversy over the balance also is due in part over failure to identify what strategic forces are intended to deter. The issue is not whether the Soviets would go to war just because they perceive that
they would suffer less damage or could recover faster than the U.S. The issue is whether such disparities will lead to more provocative Soviet behavior (or less resilient U.S. behavior) in crises. Except in very rare cases, it does not seem that even a large margin of strategic superiority can be translated into tangible political advantages by actually going to war. The large margins can be exploited politically by more risk-taking behavior in crises and it is this risk-taking behavior that must be deterred -- not decisions to go to war. For political perceptions, especially for Type II deterrence, a rough equality of counterforce and countervalue capabilities is desirable even if this requires forces far in excess of those needed for assured destruction.

This is one reason why it is unclear how much the "lose now lose later" notion really can contribute to deterrence theory. Even if the notion were completely correct and strategic forces were really incapable of deterring decisions to go to war when Soviets believe themselves to be long-run losers, there still are grounds for maintaining strategic forces for deterrence of risk-taking behavior: deterrence of decisions to go to war is not the sole purpose of strategic forces.

More importantly, the "lose now, lose later" formulation may not be quite as applicable in deterrence of strategic war as it would be for deterrence of large-scale conventional war. If the Soviets knew with high probability that a decision to go to war would mean loss of 50 percent of population and 80 percent destruction of industry and would require 10 years to recover, it is difficult to imagine a long-term future that would appear bleaker than that prospect and would therefore warrant the risk of incurring those levels of damage. The point is that the magnitude of losses being risked by a decision to go to war are an order of magnitude higher in the nuclear age than, for example, those risked by the Japanese decision to bomb Pearl Harbor: even in the worst possible case, Japan could not have expected a war outcome of such devastation as those certain in an all-out nuclear war. There is merit in attempts to persuade the Soviets that their future is bright, but this particular theory does not seem to be persuasive enough to warrant large scale reductions in deterrent forces on grounds they are irrelevant (in addition, as explained above, deterrent forces would be needed in any case to deter actions other than the decision to go to war).
IV. HOW IS UNCERTAINTY RELATED TO DETERRENCE?

Most analysts agree that uncertainty plays an important role in deterrence, but are divided on the question of whether more or less uncertainty should be sought to enhance deterrence.

On the one hand, there are those who argue that simply as a matter of fact, enormous uncertainty about the outcome of a Soviet first-strike (or whether the U.S. would launch on warning) is an important component of the deterrence of such attacks [Gray, 1973; Greenwood and Nacht, 1976]. George and Smoke provide empirical evidence that if the initiator of a prospective attack believes that the risks of the options available to him are not calculable or controllable, that usually is a sufficient condition for deterrent success (with respect to those options) [1974].

Others argue more normatively that such uncertainty is good since it contributes to stability [Lambeth, 1972], while still others view uncertainty as a key factor that prevents deterrence from failing. Steinbruner, for example, points out that to the extent that carrying out a threat is irrational, the entire strategy of deterrence is undermined; it is fairly well accepted that execution of a spasm response would in fact be irrational in the event of a Soviet first-strike, even though that is what is threatened in such contingencies [Lambeth, 1976; Leites, 1974; Rosecrance, 1972]. Steinbruner argues that if both sides were perfectly rational, the first side to realize that response would be irrational would simply preemptively attack with confidence that no response would be triggered; hence deterrence would fail. However, because there is uncertainty about whether the opponent would really behave rationally after an attack, deterrence is saved from failure [Steinbruner, 1976].

The Europeans seem to recognize that although U.S. strategic response may in fact be unlikely, uncertainty about U.S. behavior, coupled with the magnitude of the umbrella threat, is sufficient to deter; ultimately, it is the risk of escalation -- not the certainty
of retaliation -- that makes the difference for deterrence [Davis, 1976]. This same kind of reasoning is used to argue against the need for strategic options: even if a massive U.S. response would not be rational after a limited strategic attack, Soviet uncertainty about such a response is sufficient to preserve deterrence [Greenwood and Nacht, 1976]. Similarly, Russett has argued that there is no danger in moving away from assured destruction targeting to a countercombatant posture since the possibility of countercity retaliation would remain even though targeting had shifted; hence, deterrence would not be weakened even though assured destruction was less certain [Russett, 1974].

Some advocate increasing uncertainty deliberately: a comprehensive test ban, for example, is sometimes justified on grounds that it would decrease confidence in weapons capabilities and would reinforce mutual assured destruction stability since neither side could confidently mount a first strike [Scoville, 1972a]. The most ardent proponents of increasing uncertainty, however, are those who believe that risks can be manipulated in order to enhance deterrence. Rosecrance, for example, using notions originally conceived by Kahn [1962] and Schelling [1966], suggests that deterrence can be enhanced by magnifying uncertainties and letting the aggressor see the risks of escalation to high levels (without either side even wanting such escalation) [Rosecrance, 1972; 1975].

Finally, there are those who argue against certainty on grounds that perversities may arise in a world where the necessity for certainty shapes strategic arms policy. For example, Perle argues that uncertainty generally helps stabilize the strategic balance and that certain stabilizing deployments (such as mobile missiles) might be banned in SALT simply because of the uncertainties involved in verification of how many mobile missiles are deployed by each side. Similarly, too much information (i.e., certainty) generated by on-site inspection could lead to destabilizing changes in perceptions of the balance (e.g., discovering Soviet reliabilities are much lower than previously thought) [Perle, 1973]. Panofsky argues that if the risk of nuclear escalation could in fact be minimized by the controlled use of nuclear weapons (which he thinks is unlikely) then the development of that capability to control
risks would remove a factor (mentioned earlier by Rosecrance) that helps deter the outbreak of a large war: the risk of uncontrolled escalation [Panofsky, 1974].

In contrast to proponents of uncertainty, there are those who believe that reliance on uncertainty about U.S. behavior is not sufficient for a good deterrent. Instead, the certainty that a deterrent threat could and would be carried out should be enhanced by formulating concrete and specific plans for the implementation of such threats and for making options flexible so that the magnitude and nature of the threatened response was one that the enemy would believe that the U.S. really would carry out [Iklé, 1973; Schlesinger, 1974]. It is argued that if credibility of a deterrent threat is low, the effectiveness of deterrent can still be improved by increasing the magnitude of the threat; however beyond a certain point, additional destruction adds little to deterrence [Rosecrance, 1975]. Therefore, uncertainty almost has to be relied on to maintain deterrence, but some argue precisely because these uncertainties exist, it is unwise to rely solely on deterrence as a means of preventing war [Rosecrance, 1972].

Proponents of limited options have emphasized the inverse relationship that exists between the magnitude of a deterrent threat and its credibility: at the extreme, the threat of massive retaliation in response to a small provocation becomes incredible. Proponents of finite deterrence argue that "deterrent security lies not in the magnitude of the destructive force we can probably deliver but in the probability that we can deliver sufficient destructive forces" [Downey, 1976; Kahan, 1975]. However, they focus primarily on whether the U.S. can deliver a minimal level of nuclear weapons and not on whether it will deliver those weapons. Most finite deterrence advocates are satisfied to rely on Soviet uncertainty about a large U.S. response, whereas LNO advocates argue that much smaller retaliatory options must be developed in order to make U.S. response more certain [Nizze, 1976a; Schlesinger, 1974; Van Cleave and Barnett, 1974].

Included among those who prefer certainty to uncertainty are those who argue that "brinkmanship" is an unreliable means of assuring deterrence.
The "residual fear of war" [Kahn, 1962] and the "threat that leaves something to chance" [Schelling, 1966] may fail to deter because:

- The enemy may decide an escalatory bluff is less costly than giving in;
- Crisis decisionmaking is likely to distort perceptions and make misreading of signals and inappropriate response likely;
- A nation cannot back down before every irrational threat or it would have no deterrence [Kashi, 1975].

Morton Halperin suggests that U.S. deterrent posture must be credible enough that no "clever briefers" in the Soviet Union could possible argue that a Soviet first-strike would mean victory and a Soviet second-strike would mean defeat [Kahan, 1975].

A related argument favoring a high level of deterrent credibility is derived from cybernetic theory. This argument is that decisionmakers crave certainty, but they also wish to avoid value conflicts which arise when two incompatible goals are pursued simultaneously. One means of avoiding this conflict is to view the nation in a "risk war now, lose later" situation. This allows the decisionmaker to pursue a value (such as national prestige) at the risk of war on grounds that by not doing so war or defeat would certainly come later [Snyder, 1976]. Empirical evidence from Pearl Harbor and other instances of deterrence "failure" suggests that this kind of "lose now, lose later" belief underlies many decisions to go to war [Brown, 1977; Rosecrance, 1972]. The implication for deterrence from these findings seems to be that a deterrent threat must be very severe and very certain in contexts where the adversary being deterred perceives that it will lose in the long run with certainty unless the risk of war is undertaken. A perfectly invulnerable nuclear capability (along with a perception that it will be used) might impose sufficient reality constraints so that it would be difficult to perceive the situation in "lose now, lose later" terms [Snyder, 1976].
OBSERVATIONS

Whether uncertainty helps or hinders deterrence depends on the nature of the uncertainty. Certain types of uncertainty do indeed contribute to deterrence -- including uncertainty about first-strike outcomes, and uncertainty about whether the U.S. will choose to escalate a confrontation.

However, certain types of uncertainty do not contribute to deterrence. It seems unnecessarily risky, for example, to rely on Soviet uncertainty about U.S. rationality to preserve deterrence, if it can be avoided. The strongest deterrent would be one under which there was absolutely no uncertainty about U.S. ability or willingness to escalate in a crisis situation. The development of a broad range of flexible nuclear options and some clear thinking about political priorities would help bolster both the U.S. capability to escalate and willingness to do so if necessary.

Even in the polar case of a Soviet countercity first-strike, it is not clear that we need to rely on Soviet uncertainty about U.S. rationality in order to persuade ourselves that deterrence will work. If the Soviets launched an all-cities attack and the U.S. responded with a counterforce attack, who would be in a better bargaining position after that first exchange? Every subsequent Soviet move would be made with the knowledge that its cities are held hostage and that the U.S. willingness to launch against cities is no longer self-deterred by fears of a Soviet countervalue response (any marginal damage done by additional Soviet strikes against cities would be considerably smaller than that the Soviets would incur by making those strikes). So what would have been gained by such a strike? The outcome of a rational U.S. response (i.e., counterforce) would be as undesirable as the outcome of an irrational U.S. response -- so uncertainty might not be as important a factor in deterrence of these massive attacks as is supposed.

This is not to deny the increasing necessity for bluffs and risk-taking required in a world of parity or Soviet superiority. It is only
to say that minimal reliance should be placed on a deterrent posture that requires Soviet uncertainty about U.S. rationality to be the linch-pin upon which deterrent success lies.

Another point is that uncertainty on the U.S. side should not be categorically treated as unacceptable. For example, uncertainty about deployment of mobile missiles might well be worth risking if mobile missiles contribute to stability, especially if even fairly large clandestine cheating does not make a difference. So in looking at different types of uncertainty, analysts should always ask the question: does this uncertainty make a difference for deterrence? In some cases it will, in some cases it will not, and that difference ought to be reflected in policy recommendations made to deal with uncertainty.
V. HOW IS STABILITY RELATED TO DETERRENCE?

There is disagreement over whether deterrence is threatened by instability in the strategic relationship between the U.S. and Soviet Union. The most sanguine estimate is that mutual deterrence now exists and will remain stable regardless of the pace of the arms competition or state of the strategic balance (within a broad range) because a credible first-strike by either side is technologically impossible [Lambeth, 1972]. Similarly, Greenwood and Nacht note that if either side had a disarming first-strike capability, the world would be very unstable, but such a capability is not remotely possible [Greenwood and Nacht, 1976], although on grounds of arms race stability, they would prefer that U.S. counterforce improvements not be pursued.

An alternative view is that stability is quite fragile and that therefore there should be some concern about instability resulting in a failure of deterrence. Stability can be threatened by one of two incentives:

- Profit-maximizing incentives -- if the differential in outcomes between striking first and striking second is large, there may be an incentive to preempt [Perle, 1973];
- Loss-minimizing incentives -- if a nation believes the enemy has a first-strike capability, there is a temptation to move to a launch-on-warning posture to avoid loss of strategic forces [Carter, 1974; Iklé, 1973]; but this increases the likelihood of an accidental or miscalculated launch.

However, the belief that mutual deterrence is fragile leads analysts to opposite conclusions about what to do about it. The majority of analysts argue that since there is no escape from a MAD posture (since defense or damage limitation is too costly given the size of Soviet forces [Rosecrance, 1972]), the best strategy is to minimize the likelihood of instability arising. On these grounds, certain analysts attack
any U.S. attempts to improve its counterforce capabilities as provocative and destabilizing; within this group, there are those who believe that no distinction can be made between weapons intended for first-strike and those intended for second-strike -- therefore the U.S. should eschew all efforts to improve counterforce to avoid crisis instability [Panofsky, 1974; Carter, 1974]. An additional view is that improved counterforce could be destabilizing in the arms race sense, so that U.S. should engage only in restrained counterforce improvements [Greenwood and Nacht, 1974].

Similarly, there is an argument against widening the range of U.S. limited options (apart from improving counterforce) on grounds that more options make nuclear war more feasible and ultimately more likely [Panofsky, 1974]; this is generally coupled with the belief that once the nuclear firebreak is crossed, escalation control is very difficult [Downey, 1976; Panofsky, 1974; Scoville, 1974]. Therefore, on grounds of escalation stability, options should not be pursued and the threshold level of nuclear use should remain high.

Aside from such restraints on U.S. behavior, some advocates of maintaining stability argue for a broad program of research and development to ensure that the U.S. can maintain the current balance and stability in the face of Soviet force improvements that challenge mutual deterrence [Pfaltzgraff, 1975; Steinbruner, 1976]. This kind of approach appears to be addressed to the question of dynamic stability -- that is, how fast factors contributing to static stability can change over time [Perle, 1973]. Presumably, policies to deal with dynamic stability could be logically advocated by both those who believe static stability exists and those who believe that static stability is fragile.

Another major line of argument is that because mutual assured destruction may not be stable, it should be modified or augmented with limited nuclear options to provide for escalation control in the event that deterrence fails for one reason or another. Iklé argues that if assured destruction capabilities are needed just for deterrence, and not warfighting, there is no reason that a rapid response is necessary. Therefore, one way of reducing the instabilities in the current system
would be to make retaliatory systems more survivable and relax the requirement on swift responses [Iklé, 1973]. Some advocates of mutual assured destruction agree that speedy retaliation is unnecessary, but argue that the U.S. should continue to rely on assured destruction (even if slow) rather than switching to counterforce targeting or damage limitation since these alternatives are believed to threaten crisis stability [Cohen, 1973; Panofsky, 1974].

Others argue that with flexible options, the chances for escalation control are much better than in a pure mutual assured destruction world, and therefore intrawar deterrence can prevent accidents, miscalculations or other unintentional small attacks from becoming all-out holocausts [Schlesinger, 1974; Van Cleave and Barnett, 1974]. Proponents of this view note a paradox in the views of those who simultaneously argue that mutual assured destruction is very stable (i.e., large changes in the balance will not change the mutual hostage relationship between the U.S. and Soviets), but that flexible options will be destabilizing [Panofsky, 1974]. If the balance really is stable, the addition of flexibility would make no difference in any case, say proponents of LNOs, but if it is not stable, then something must be done about growing asymmetries favoring the Soviets -- and this should include the development of limited nuclear options [Schlesinger, 1974; Van Cleave and Barnett, 1974].

In addition to options, some believe that in order to limit Soviet profit-maximizing incentives, the U.S. should improve its missile accuracies and increase yields in order to minimize disparities in the pre- and post-attack ratios of U.S. and Soviet counterforce potential [Nitez, 1976a]. This is refuted by opponents of improved counterforce on grounds that while this might minimize the difference between the value of a first-strike and the value of a second-strike in terms of the military balance, it also provides preemption incentives that would lead to crisis instability.

Finally there are those who believe that even if mutual assured destruction is thought to be stable, it cannot eliminate the likelihood of nuclear war; therefore, targeting cities is immoral and should be abandoned. Proponents of this view argue that the direction of strategic
procurement should be in favor of more defense and more damage limiting capabilities [Brennan, 1972; Bailey, 1972] or a shift to a different kind of targeting [Iklé, 1973; Russett, 1972].

OBSERVATIONS

In most discussions of stability, analysts do not clearly distinguish between arms race stability, crisis stability and escalation stability. Those who believe SALT should be aimed at improving stability appear to believe that SALT can simultaneously contribute to all three types of stability -- yet it is not clear that this is true. The mobile missile, for example, poses tremendous arms control and verification problems and is generally disparaged on grounds that mobile deployments would touch off a destabilizing arms race. Two points need to be made:

- It has never been clearly demonstrated how arms race instability might affect deterrence; on budgetary grounds, arms races might be reprehensible, but no clear case has been made that arms race instability directly endangers deterrence; are the Soviets really more likely to launch an attack because they have been arms racing? Some plausible connections can be imagined: if the Soviets believe they are going to lose a race in the long run they may preemptively strike while they still have an advantage; alternatively, the general historical flavor of the U.S.-Soviet relationship is likely to be more hostile prior to a crisis if an arms race is underway, and hence decisions to attack or not may be affected.

- However, there is a trade-off between arms race stability and crisis stability in the mobile missile case, since the invulnerability of mobile missiles could substantially alleviate the crisis stability that would otherwise arise once both sides attain mutually devastating counterforce capabilities; are those who oppose mobile missiles really certain they wish to trade a higher probability of crisis stability for a lower probability of arms race stability, given that the connections between crisis stability and deterrence appear much stronger
than the relationship between arms race stability and deterrence?

Similarly, there is a trade-off between crisis stability and escalation stability. The views of both sides on the issue of whether flexibility is desirable will be presented in detail in the next section. Briefly, those who oppose more flexibility argue that it will degrade deterrence by making use of nuclear weapons more likely; those who favor flexibility are willing to concede that this might increase the probability of war, but that this will be more than offset by a higher probability of intrawar deterrence which will keep war limited.

The point of both of these examples is that the relationship between deterrence and "stability" is not straightforward, since it depends on what kind of stability is being considered. In order to assess the net impact of a policy change on the effectiveness of deterrence, the impact of that change on all three types of stability must first be considered and then the interactive effects of all these changes in stability must be related to deterrence. The current strategic literature often fails to reflect such nuances.
VI. HOW IS FLEXIBILITY RELATED TO DETERRENCE?

There seem to be three major positions on the issue of flexibility and deterrence:

- More flexibility is good because it enhances deterrence credibility; in addition, it improves intrawar deterrence in the event deterrence fails and therefore helps keep nuclear wars limited and damage limited; flexibility requires improved U.S. counterforce capabilities;
- More flexibility is good for the above reasons, but it does not require improved U.S. counterforce capabilities;
- More flexibility degrades deterrence by making the prospective costs of aggression lower and by making U.S. first use more likely because the consequences of going nuclear appear lower; in fact, escalation is very hard to control, so any limited exchange will almost inevitably become an all-out exchange; the U.S. therefore has sufficient flexibility in its forces.

The primary rationale for flexibility is that deterrence of an initial attack alone is not enough; nuclear exchanges may begin without a calculated first-strike, through accidents, unauthorized launches or by third-party deception [Iklé, 1973; Schlesinger, 1974]. In the event that nuclear exchanges begin in such a fashion, the capability to respond in a limited fashion would minimize the likelihood that a nuclear holocaust would result -- intrawar deterrence would be achieved by restrained responses that invited response in kind and by withholding an assured destruction reserve that would minimize enemy incentives to retaliate against U.S. cities [Lambeth, 1976; Russett, 1974; Schlesinger, 1974].

A second important rationale for flexibility was that a threat of massive damage such as that offered by assured destruction was not credible in the face of limited Soviet attacks on military targets or cities. A related argument was that assured destruction was not a
credible threat for limited or even massive attacks against Europe since the Soviets might believe that the U.S. interests in Europe would not be strong enough to warrant risking an all-cities response by the Soviets -- which is what an assured destruction retaliation would invite [Martin, 1974]. This problem of "self-deterrence" [Lambeth, 1976], also referred to as "deterring our deterrent" [Nitzze, 1976a], could be overcome if options were available. Thus, flexibility offers a means of resolving the chronic dilemma between ex ante threats and ex post incentives by threatening limited responses that are in fact rational to carry out [Davis, 1976; Nitzze, 1976a; Rosecrance, 1975].

Aside from primary deterrence, limited options are thought to be useful for demonstrating resolve after conflict begins; this strengthens intrawar deterrence by demonstrating a willingness to escalate if necessary to achieve objectives [Schlesinger, 1974]. George and Smoke argue that in limited war scenarios, the capabilities of each side are not in doubt (although Schlesinger would argue that prior to 1974, U.S. limited war capabilities were in doubt), but intentions are in doubt in such situations [George and Smoke, 1974]. This suggests that measures taken to clarify intentions can strengthen intrawar deterrence (whether limited options in fact do clarify intentions is not addressed by George and Smoke).

The issue of how to obtain enough flexibility to improve deterrence divides those who advocate more flexibility. Schlesinger argued that a capability to respond in kind to any type or level of attack was important to deterrence and his statement that "we do not intend that the Soviet Union should have a wider range of options than we do" [1974] suggests that a balance of strategic options is required for deterrence. This notion provides part of the justification for arguing that higher accuracy was needed for U.S. forces; in addition, higher accuracy was seen as minimizing collateral damage, thereby improving the likelihood that the Soviets would perceive U.S. responses as really being limited. A third rationale for increasing accuracy in order to gain flexibility was to improve the efficiency of U.S. hard target kill capabilities in order to threaten "specialized sets of targets (possibly of concern to allies)" [Schlesinger, 1974].
However, others question the need for a capability to respond in kind and believe that flexibility can be achieved by better command and control and changes in targeting capabilities; improved counterforce capabilities would be destabilizing [Carter, 1974; Davis, 1976; Greenwood and Nacht, 1974]. A concomitant benefit of improved command and control would be a reduction in the likelihood of accidents [Scoville, 1974]. Some of these critics of improved counterforce acknowledge that there would be some degradation in deterrence due to a lack of symmetry in options, but argue that this would be outweighed by the gains in stability of not provoking the Soviet Union with improved counterforce capabilities [Kahan, 1975].

Proponents of flexibility who oppose improved counterforce argue that the primary purpose of flexibility should be to try to limit conflict at the lowest possible level of violence [Davis, 1976]. The primary objective in attacking counterforce targets is to demonstrate resolve -- and for that, relatively small counterforce capabilities are needed. Any LNOs designed to impose significant military costs would have to be quite large (i.e., hundreds of weapons), and flexibility advocates who oppose improved counterforce believe that it is ill-conceived to try and impose substantial military costs since appreciable damage limitation is simply not possible except at provocatively high levels of counterforce [Davis, 1976].

Another argument against improving counterforce is that response in kind does not require that the U.S. respond to attacks on silos with retaliation against Soviet silos: for purposes of demonstrating resolve, many other military targets exist which do not require a hard target kill capability and destruction of these would be equally effective for signalling intentions [Davis, 1976; Downey, 1976; Scoville, 1974]. Even if symmetrical response in kind were necessary, the U.S. would have sufficient Minuteman ICBMs to carry out the retaliation -- unless the Soviet first-strike were overwhelming, in which case improved accuracy on the residual U.S. forces simply would not add much capability; improved accuracy on SLBMs would provide little damage limitation in this scenario since the Soviets would still maintain their entire SLBM force
to use on U.S. cities even if the U.S. knocked out all remaining
Soviet ICBMs [Scoville, 1974].

In response to those who claim that flexibility would increase the
likelihood of war, some argue that flexibility is a matter of degree and
that the U.S. always has had flexibility: a small addition of flexi-
bility to forces will increase credibility, not impair deterrence [Van
Cleave, 1974]. Advocates of increased counterforce for limited options
say that a disarming first-strike by either side is not technically
feasible; hence, there should be no reason for marginal improvements
in accuracy to be provocative or destabilizing [Schlesinger, 1974].
Others argue that there is no way of telling what actions will weaken
the deterrent effect since no one knows what actions in fact have a
deterrent effect; arguments that improvements in U.S. warfighting capa-
bilities will impair deterrence are "probably fallacious" [Gray, 1973].

Finally, there are those who claim that the net effect of flexibility
will be to improve deterrence compared to an assured destruction-only
threat; however, paradoxically, even with flexibility the deterrent
threat might not be high enough since the product of credibility (which
has been enhanced) and magnitude of the threat (which has been reduced)
might not be large enough to deter [Rosecrance, 1975].

There are three major lines of argument against flexibility. The
first is that flexibility could decrease deterrence and make nuclear
war more likely in the following ways:

- If the Soviets believed that U.S. responses would be limited,
  they might be more prone to risk-taking behavior [George and
  Smoke, 1974; Kahan, 1975; Scoville, 1974]; for irrational
  leaders in particular, deterrence would be more effective
  if there were fewer opportunities available in which they
  might believe they could fight a nuclear war and win [Scoville,
  1974];

- If U.S. leaders believed that "clean" limited wars were pos-
  sible, the political costs of initiating nuclear war will be
  lower, so they will be more prone to use nuclear weapons in
  crisis [Kahan, 1975; Panofsky, 1974; Tsipis, 1975b]. Stein-
  bruner claims that cybernetic theory indicates that making
options available increases the likelihood they will be used [Steinbruner, 1976];

- The command and control measures required for flexible responses, the increased use of alerts and the "playing around" with weapons under flexible options makes accidental launch more likely and this would offset any improvements in intrawar deterrence provided by flexibility [Ball, 1974].

The second major argument against flexibility concerns the feasibility of limited nuclear war. A large number of analysts are skeptical that escalation can be controlled once the nuclear threshold has been crossed [Ball, 1974]. There appear to be five factors which make escalation control in a nuclear war questionable:

- There will be inevitable collateral damage, even with surgical strikes [Scoville, 1974]; the intended consequences of limited strikes are likely to lead to misperceptions and unintended escalatory moves [Kahan, 1975; Steinbruner, 1976].

- Steinbruner argues that cybernetic theory suggests that the Soviets would be predisposed to view even limited attacks as mortal blows and this would provoke outrage that would heighten escalation [1976]; similarly, Downey argues that both sides would want to inflict equal pain on each other and would "leap-frog each other briskly up to Armageddon" [Downey, 1976];

- Limited war requires two to tango; there is enormous uncertainty about whether the Soviets would remain restrained under limited attacks [Carter, 1974; Drell and von Hippel, 1976; Steinbruner, 1976]. Some argue that Soviet doctrine is strongly opposed to limited war, and that even if the Soviets wanted to engage in limited options, technical constraints might force them to escalate for lack of sufficient options [Kahan, 1975];

- It is not clear that the U.S. political system would allow a President to engage in limited nuclear options, because the public pressure for revenge would be too great [Cohen, 1973];
If hard target kill capabilities exist on both sides, there is no incentive to withhold forces after a limited exchange since each side will assume that the other will attack them [Kahan, 1975]; therefore any limited exchange will escalate.

A third argument concerning flexibility addresses the question of whether improved U.S. counterforce capabilities would improve deterrence. Opponents of improved counterforce argue that since there is no difference between counterforce capabilities intended for a first-strike and those intended for a second strike, three Soviet reactions are possible:

- A move to mobile missiles, which would hurt arms control and in the long run might affect deterrence [Tsipis, 1975a];
- A launch on warning posture to avoid loss of strategic forces [Kahan, 1975; Scoville, 1974; Tsipis, 1975a]; since the Soviets have nearly three-quarters of their strategic capability in land-based missiles, this reaction is quite possible;
- Preemptive attack in crisis to avoid loss of strategic forces (Kahan, 1975).

This argument against improved counterforce is held even by some analysts who claim that a disarming first-strike by either side is technically impossible; they argue that stability requires that "there should be neither a real nor a 'perceived' vulnerability of any major component of the strategic deterrent forces on either side" [Drell and von Hippel, 1976]. Therefore, even though a U.S. disarming first-strike would not be feasible, improved counterforce would appear like a push for a first-strike and would erode the mutual deterrent [Scoville, 1974].

**Observations**

The argument of the opponents of increased flexibility can be put in the form of the following syllogism:
1. LNOs increase the probability of war;
2. But once war begins, it inevitably escalates to an all-out level;
3. Therefore it is preferable to rely on uncertainty about an assured destruction response rather than increase the probability of all-out war.

There are two grounds for questioning the validity of assertion #1. From the U.S. perspective, the argument is that the availability of options makes their first use more probable; yet historical experience does not bear this out. A decision to use tactical nuclear weapons in an already-raging conventional war is far less momentous than the decision to enter strategic nuclear war; yet even though the U.S. had tactical nuclear options in Korea and Vietnam, it failed to use them. Granted, there were unique factors present in both cases which led to this restraint, but the point is that the U.S. was restrained despite the presence of options. It is not clear that the presence of strategic options would appreciably increase U.S. willingness to initiate the use of nuclear weapons -- especially given that U.S. cities are hostage to any Soviet retaliatory strikes (a factor not even present in the decisions about whether to use tactical nuclear weapons). The availability of options would increase the probability of U.S. response after a limited Soviet attack, but that is precisely the justification for moving to LNOs, since this increased probability bolsters deterrence.

From the Soviet perspective, it is alleged that their willingness to initiate limited wars will increase if they know that the U.S. response will be limited. But arguing from the premises of the opponents of LNOs, this would not make sense. If all limited wars inevitably become all-out wars, then a U.S. LNO response is as fearsome as an all-out response since the LNO response eventually will trigger the all-out response in any case. (This argument is invalid if one claims that the Soviets may believe that limited wars are possible; however, such a contention would be difficult to support using Soviet doctrinal writings -- see [Lambeth, 1976].) But even if the Soviets did believe in limited wars, why would not the residual uncertainty about whether
the U.S. would move to assured destruction be equally effective in an LNO world as flexibility opponents claim it is currently? If this uncertainty about assured destruction is reliable enough to serve as a deterrent in a world without improved flexibility, why can it not continue to do so even after flexibility is augmented?

The issue of improved counterforce has been addressed at various points in this review. Suffice it to say, there are many reasons to believe that improved counterforce will increase the risks of crisis instability and arms race stability in the future. In the absence of a compelling reason to move towards improved accuracy and yield, it is difficult to justify incurring such risks. It does not appear that accuracy really buys any militarily significant additional capability nor does it expand flexibility beyond what is already available in current forces; hence, the burden of proof still lies with advocates of improved counterforce to demonstrate why the risks of moving in that direction are worth taking.
VII. HOW CAN SALT CONTRIBUTE TO DETERRENCE?

Most analysts believe that SALT and stability go together — that the primary goal of SALT (if not the outcome) has been to stabilize the strategic balance and cast mutual assured destruction in concrete [Newhouse, 1973]. Arms race stability and crisis stability are the two factors on which SALT is seen to impact the most, while escalation stability is usually ignored since it only becomes relevant after primary deterrence has failed. Most of the policy recommendations designed to do something about deterrence fall into one of three categories: (1) Preserving mutual assured destruction; (2) Limiting the strategic arms competition; (3) Maintaining the strategic balance. Before proceeding with a review of these areas, the point of view of skeptics about SALT should be outlined.

Those who believe that stability is no real problem anyway (since neither side can obtain a disarming first-strike capability), argue that on strategic grounds, there is no reason to engage in SALT: deterrence will be maintained in any case [Lambeth, 1972]; however, there might be other nonstrategic reasons to continue the SALT dialogue. Those most upset with the current trends in SALT are those who believe that mutual assured destruction is disastrous. In their opinion, SALT should have been used to limit offensive arms so that effective defenses could be erected and damage limitation could be pursued [Brennan, 1972b]. In their opinion, no SALT at all would have been preferable to the agreement that was reached in 1972.

PRESERVING MUTUAL ASSURED DESTRUCTION

There are three approaches to maintaining mutual assured destruction:

- Limit improvements that would make retaliatory forces vulnerable;
- Reduce the number of vulnerable retaliatory forces or improve their invulnerability;
- Limit improvements that would make population less vulnerable.
Among those who see SALT as an avenue to crisis stability, the first approach has been the most heavily emphasized. For the land-based forces, four means of limiting counterforce improvements have been suggested for keeping land-based missiles invulnerable:

- Limits on numbers of heavy missiles;
- Limits on the size of ICBMs;
- Limits on the yield (or number of MIRVs) per ICBM;
- Limits on the accuracy of ICBMs or SLBMs.

Limits on the number and size of heavy ICBMs were included in the SALT I agreement. Some have pushed for a ban on all MIRV deployments, since without MIRVs, the counterforce problem is much less difficult [Carter, 1974; Coffey, 1973]. Others argue that limits on heavy ICBMs are irrelevant since in the long run, the real threat to U.S. strategic forces will come from MARVed SLBMs [Downey, 1976]. Nitze argues that limiting throw-weight to 200 pounds per ICBM would eliminate the counterforce problem, but it is not verifiable [1974]. Most analysts acknowledge that accuracy limits would be similarly unverifiable, so an alternative approach is to establish test limits in order to slow down the pace of improvements and reduce confidence in accuracy enough that a first-strike would be made too risky [Scoville, 1972; Tsipis, 1975a]. Others argue for a comprehensive test ban on similar grounds [Scoville, 1972]. However, this view is challenged by those who believe that accuracy improvements are inevitable since systems such as Global Positioning Satellites will be pursued for other reasons (i.e., navigation); in addition, accuracy improvements may ultimately be stabilizing to the extent that they push Soviet forces out to sea [Lodal, 1976]. Some of those who do regard both accuracy and yield improvements as destabilizing acknowledge that both pose difficult SALT verification problems and believe that unilateral restraint on U.S. improvements in these areas is the best approach [Carter, 1974].

Some claim that the vulnerability of bombers could be reduced by a SALT ban on close-in SSBN operations [Downey, 1976]; many analysts believe that restrictions on ASW to maintain the invulnerability of SLBMs would be equally desirable. Such restrictions could include limits on numbers of attack submarines [Scoville, 1972a]; a ban on continuous
tracking of SSBNs [Kaplan, 1975]; prohibitions on the installation of large sonar arrays for ASW detection [Kaplan, 1975]; or the establishment of protected enclaves for each side’s SSBN fleet in which attack submarines would be prohibited [Downey, 1976; Kaplan, 1975; Scoville, 1972a].

Another approach to the counterforce problem would be to gradually phase out vulnerable ICBMs on both sides [Carter, 1974; Kahan, 1975; Quanbeck and Blechman, 1973]. Some believe that mobile ICBMs should be deployed to reduce ICBM vulnerabilities [Cohen, 1973; Nitze, 1976b], but others worry that this would cause substantial verification problems at SALT [Kahan, 1975; Lodal, 1976]. In addition, some have argued that mobile missiles do not eliminate the vulnerability problem since they too will be vulnerable once MARV accuracies are achieved in the late 1980s [Downey, 1976]. An alternative approach would be to permit ICBM-only ABM systems to protect vulnerable ICBMs [Lodal, 1976].

Defense advocates argue in the complete opposite direction and say that ABM should be permitted for population defense and for deterrent forces, but not for any other military forces [Iklé, 1973]. Others say ABM deployments would be useful only if MIRVs were limited [Cohen, 1973]. However, many regard ABM deployments as destabilizing since they threaten the retaliatory capability of the other side [Perle, 1973]; in addition, some have noted that limits on ABMs ensure that limited nuclear options can be carried out, whereas if large ABM deployments are permitted, retaliatory strikes would have to be very large to ensure that even some missiles get through [Kahan, 1975; Panofsky, 1974; Scoville, 1974].

LIMITING THE STRATEGIC ARMS COMPETITION

There are two avenues that can be used in SALT to limit the strategic arms competition:

- Reduce launcher and MIRV ceilings;
- Limit new systems or improvements on old systems.

Many analysts have proposed that some sort of phased ceiling reductions be pursued in order to attain a stable strategic balance at lower
levels of forces [Nitze, 1974; Pfaltzgraff, 1973; Scoville, 1972]. Some have argued that ceiling reductions of as high as 50–75 percent would be worthwhile and that even at such low ceiling levels, the problem of cheating would not create any significant instabilities [Rathjens, 1974]. Others argue, however, that the risks of SALT violations become much greater as ceilings drop and certain uncertainties in weapons parameters are no longer tolerable since there is not a margin of extra forces to hedge against these uncertainties [Perle, 1973]. Some of those who have proposed alternatives to mutual assured destruction do not favor ceilings since they would degrade the capability to perform other kinds of targeting [Russett, 1972]. But even advocates of mutual assured destruction sometimes argue that ceiling reductions would be meaningless if systems such as cruise missiles were not included and could be deployed without limit [Vershbow, 1976].

It is believed by some that limits on the production of new systems can help attenuate the arms race [Aaron, 1974]. In pursuit of that objective, some have advocated banning mobile missiles [Lodal, 1976]; cruise missiles [Tsipis, 1975b; Vershbow, 1976]; and ABM or other defensive systems thought to accelerate the arms race [York, 1973]. In contrast, others argue that the verification problems associated with mobile missiles are irrelevant since ceilings are high enough that cheating would not make any difference in the balance [Nitze, 1974]; in this view, the stabilizing features of mobile missiles are thought to outweigh arms race stability problems. Similarly, there are those who believe cruise missiles could be quite stabilizing since they increase second-strike retaliatory capabilities without posing a first-strike threat [Lodal, 1976].

Another approach is to limit testing on new systems; some advocate this approach to limiting MARVs since once MARVs have been tested, there would be tremendous problems in verifying the number deployed [Downey, 1976].

Although some believe that abandonment of SALT -- even for a short time -- would extend the arms competition and reduce security for both sides [Lodal, 1976], others have argued that the U.S. bargaining chip
approach to SALT contributes to the arms race and is an unwise policy [Greenwood and Nacht, 1974]; some even argue that SALT is too expensive a luxury if it requires continued development and deployment of new systems as bargaining chips and that unless this approach is abandoned, SALT should be abandoned [Warnke, 1975].

MAINTAINING THE STRATEGIC BALANCE

Much of the current debate about the balance has already been reviewed in a previous section. This section will simply quickly review the opinions of those who have specifically tied their concerns about the balance to a proposed SALT negotiation goal. Some of those who worry about the balance argue that SALT I gave the Soviets a vast potential for superiority that may lead to "Finlandizing" among allies if SALT II does not correct the imbalances [Brennan, 1972]; others are more sanguine and believe that SALT I will not make the Soviets any more eager to escalate in a crisis, but may make them more resistant to U.S. efforts to weaken Soviet influence in the Middle East, Europe and Asia; so long as NATO remains strong, Finlandization will not be a problem. However, without numerical equality in launchers under SALT II, the U.S. will be unnecessarily risking escalation in future crises [Kaplan, 1975]. Others have argued for equal ceilings [Kahan, 1975; Nitze, 1974], on grounds that while equal aggregates may not be militarily significant, they are politically significant [Lodal, 1976]. More comprehensive approaches to the balance argue for U.S. avoidance of inferiority in launchers and an agreement that will not permit the Soviets to obtain a lead in both warheads and payload simultaneously [Kahan, 1975]; an alternative would be to pursue roughly equal ceilings on the throw-weight of MIRVed missiles and roughly equal ceilings on aggregate MIRVed and unMIRVed launcher throw-weight (taking bombers into account) [Nitze, 1974].

OBSERVATIONS

Many analysts appear overly preoccupied with narrowly focused concerns about crisis stability or arms race stability and as a consequence may be advocating a future world for the U.S. that would not really be
very desirable. For example, the following might be considered:

- Many appear to think that ceiling reductions would be very beneficial since these would help limit the arms race and perhaps improve stability; such outcomes ignore consideration of what vast reductions in ceilings might do for Type II deterrence. In particular, consider the world that Rathjens advocates in which there has been a 75 percent reduction in ceilings and both sides have 600 launchers; if one considers Soviet bomber defenses, future counterforce capabilities, breakout potential (such as converting SS-20s to SS-16s), civil defense and inherent advantages in clandestine deployments, how credible is a U.S. nuclear umbrella? Large ceilings at least provide a hedge against this kind of world, yet the impact of ceiling reductions on Type II deterrence is rarely addressed;

- Similarly, who has a comparative advantage in a world of test limits and a comprehensive test ban? How does our decreased confidence in the capabilities of weapons interact with proposed ceiling reductions? Do we really want to enter a world in which both sides have a small number of launchers and a large amount of uncertainty about whether they work?

- As mentioned previously, the failure to consider trade-offs between arms race stability and crisis stability can lead to some perhaps undesirable SALT negotiating positions; aside from cost and arms race concerns, should not the stabilizing merits of mobile missiles be given more attention? Similarly, ABM deployments may well be destabilizing if pursued in large numbers; but might not a thin population defense considerably reduce current fears of accidental or unauthorized launches? Presumably, such a system would not pose major verification problems, yet it is hardly considered because of a prevalent belief that ABMs are not cost-effective or would be destabilizing. Some ABM systems are, but others are not.
There is no point in summarizing here what has already been a summary of a large body of literature on deterrence. The basic conclusion of this review is that more analysis of trade-offs is warranted in the debate about deterrence. What exactly do we wish strategic nuclear forces to deter, and how does the procurement of capabilities to deter one kind of action constrain or enhance our capability to deter other kinds of actions? Are strategic nuclear weapons intended only to deter decisions to go to war, or are they also intended to deter attempts at coercion? How much difference does uncertainty make for deterrence and what are the trade-offs between uncertainty and stability that necessitate policy choice? What kind of stability is most important for deterrence and how do trade-offs between arms race stability, crisis stability and escalation stability affect the deterrent forces we choose? Does flexibility matter for deterrence, and if so, how much crisis stability are we willing to give up for it? Can SALT be used to get better deterrence capabilities, or is SALT a long run threat to deterrence? How does pursuit of Type I deterrence at SALT affect Type II deterrence capabilities?

In short, there are many more questions raised by this review than answered. However, these are the kinds of questions that must be asked and answered by those whose policy decisions will be affecting U.S. deterrent capabilities as we move into the 21st century.


Downey, Thomas J. (1976), "How to Avoid Monad -- And Disaster," Foreign Policy, 16 (Fall, 1976). pp. 172-201.


