SOME THOUGHTS ON DETERRENCE

Alan Vick

October 1983

P-6916

SOME THOUGHTS ON DETERRENCE

Alan Vick

October 1983

SOME THOUGHTS ON DETERRENCE

INTRODUCTION

Deterrence theory is widely used and accepted by many American government officials and international relations scholars. The concept is quite old, appearing as early as Thucydides' *History of the Peloponnesian War*, wherein a representative of Corcyra, who is lobbying for Athenian support, argues that through an alliance with Corcyra "you will certainly become stronger, and...this fact will make your enemies think twice before attacking you."[1] Despite this early use, between 400 B.C. and 1945 deterrence received little attention outside of jurisprudence.

The use of atomic weapons against the Japanese cities of Hiroshima and Nagasaki convinced many statesmen, soldiers and scholars that, if civilization were to survive, the prevention of war must become the primary goal of the state. Bernard Brodie argued:

> Thus far the chief purpose of our military establishment has been to win wars. From now on its chief purpose must be to avert them. It can have almost no other useful purpose.[2]

This desire to prevent a nuclear war, however, was unable to overcome the collective unwillingness to disarm. As implied by Brodie, perhaps the answer lay in the deterrence of war through the ability to retaliate against any aggressor with terrible swiftness and destructiveness. Nuclear deterrence, with its stress on war prevention, offered a realistic compromise between total disarmament and business-as-usual.

Deterrence theory did not become widely accepted until the late fifties when Brodie's *Strategy in the Missile Age* and Thomas C. Schelling's *The Strategy of Conflict* appeared. Additionally, many Rand

---

[1]Thucydides, *History of the Peloponnesian War*, New York, Penguin Edition, 1972.

[2]Bernard Brodie, *The Absolute Weapon*, New York, Harcourt and Brace, 1946.

Corporation researchers contributed to early deterrence theory. Of particular interest were Wohlstetter's strategic bomber-basing study and Marshall's and Kaufman's "no-cities" or "counterforce" strategy. Wohlstetter became much more visible when his "The Delicate Balance of Terror" was published in *Foreign Affairs* in 1959. Deterrence quickly became the conventional wisdom in both the field of international relations and among the foreign policy community. Ironically, deterrence theory became widely accepted despite the lack of empirical grounding for the model. In part, its appeal was due to the popularity of Realism (the theory that politics is governed by objective laws) and the Cold War atmosphere.

Deterrence theory, because it explained seemingly perverse behavior by states and because of its precision, helped address many criticisms of Realism just as Lenin's contribution on imperialism strengthened Marxist theory. Furthermore, while the basic propositions of Realism did not lead to further interesting research, the basic propositions of deterrence generated countless new propositions and were easily applied in diverse circumstances.[3]

This paper offers a brief review and discussion of the fundamentals of deterrence theory. It offers an alternative and, it is argued, more realistic method for calculating expected value.

MANIPULATING THE EXPECTED VALUE OF NUCLEAR ATTACK

Nuclear deterrence theory is most easily understood as the mutual manipulation by threat of two states. National decision-making under these conditions is interdependent. Thus, restraint is reciprocated by restraint and attack by immediate retaliation. A potential attacker will be deterred as long as the expected value from an attack is less than the expected value from not attacking. For illustrative purposes, expected value calculations, combining both expected probability and expected utility (benefit) will be used. While helpful in presenting the basics of deterrence, there is little evidence to suggest that these calculations are representative of real-world strategic decision-making. Expected-value calculations demonstrate what a leader should think about

---

[3]Robert Jervis, "Deterrence Theory Revisited," *World Politics*, Vol. 31, pp. 289-324.

and, if he wishes to maximize expected value, what conclusions he should draw. They are not explicitly used by strategic decisionmakers in either the Soviet Union or United States.

To illustrate expected value calculations, consider the following decision tree for one country deciding whether to attack a country which may or may not retaliate. The potential attacker establishes certain estimates about the likelihood of particular responses (subjective probability) and the expected benefit from particular outcomes (utility). Combining probabilities and benefits yields expected value, as shown in the following hypothetical example:

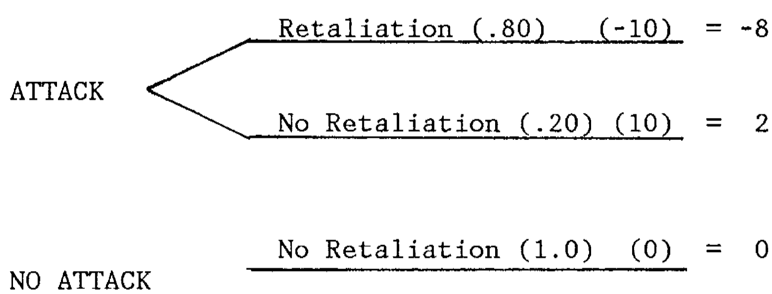

Fig. 1.1 -- Attacker's Decision Tree

where (.80) and (.20) are the respective subjective probabilities of retaliation. -10, 10 and 0 are the utilities received from the different outcomes. -8, 2 and 0 are intermediate values added together to derive expected value.

The country's expected value from an attack is:

$$(.80)\ (-10) + (.20)\ (10) = -8 + 2 = -6$$

and from not attacking is:

$$(1.0)\ (0) = 0$$

In this case, deterrence succeeds because the expected value from not attacking (0) is greater than the expected value from attacking (-6).

The deterrer will seek to manipulate the attacker's calculus such that the expected value of an attack is significantly less than the expected value of not attacking. The difference between these is equal to zero when the potential attacker is indifferent. Deterrence is in danger of failing at this point.

Expected value can be manipulated either by influencing a potential aggressor's utility function or subjective probabilities. Deterrence theorists who stress the unacceptably high level of destruction that would result from nuclear aggression are suggesting that an extraordinarily negative expected value will result for the recipient of nuclear retaliation. They attempt to convince the potential aggressor that the rewards from nuclear war are low indeed. The deterrer can do this through pronouncements, deployment of weapons or development of a targeting strategy, all threatening harm to the opponent's most treasured resources. Alternatively, the deterrer might attempt to change the very nature of the opponent's utility function by changing that society's and leader's attitudes regarding the positive payoff attached to nuclear victory. This latter course may be possible through the dialogue of strategic arms negotiations.

Theorists who stress the operational level of nuclear war and are most concerned with actual exchange scenarios tend to be more concerned with manipulating expected value through subjective probabilities These analysts seek to convince the opponent that there are no conditions under which a punishing retaliation can be escaped. They advocate a wide variety of targeting plans and deployment of weapons which can attack all types of targets in both a limited and unlimited fashion so that the opponent recognizes that retaliation will always be the most desirable option for their side and that aggression will always be followed by retaliation.

Given certain subjective probabilities and the attacker's utility from an attack followed by no retaliation and from peace, one can solve:

$$(.10)\ (x) + (.90)\ (20) = 0$$

where (.10) is the probability of retaliation, x is the utility for the attacker if retaliation takes place, (.90) is the probability of no retaliation and (20) is the utility in case of no retaliation. The solution to the equation is x = -180. The deterrer ensuring that the attacker's cost (negative utility) of retaliation is higher than or equal to the absolute value of -180 (180) will produce an expected value from attack of zero or less. This could be done by strategy or targeting changes or the use of higher yield or "dirty" weapons.

Alternatively, one can solve for probability as follows:

$$\begin{aligned}(-10)p + 10(1-p) &= 0\\ 10 &= 20p\\ p &= .50\end{aligned}$$

Given these utilities, the attacker's subjective probability of retaliation need only be (.50) to deter an attack.

While such equations help one understand the fundamentals of deterrence, they can mislead. Assigning numerical values to non-financial expected values is a highly suspect exercise. Even the illustrative uses one sees in game theoretic analyses of deterrence are risky. While this model may be adequate for some purposes, there is a tendency to forget that the numbers inserted by the theorist are truly arbitrary. The very use of such numbers implies a precision, a scientific element that is impossible to achieve in practice.

THE USE OF BOUNDARY CONDITIONS IN EXPECTED VALUE CALCULATIONS

A more realistic expected value calculation can be generalized as follows:

$$\text{expected value} = u_1p_1 + u_2p_2 + u_3p_3 + \ldots + u_np_n \qquad (1.1)$$

where:

u_n = utility from response n
p_n = probability of response n occurring

In this case the victim has an indefinite and large number of options including immediate surrender or retaliation against cities, ports, military installations, or strategic forces with any number and type of weapons. Thus, an unmanageably large number of outcomes would have to be analyzed to arrive at an expected value for attack.

Real-world national leaders have neither the inclination, time, nor mathematical skills to do the truly difficult calculations associated with real-world deterrence. Furthermore, there is no reason to believe that the opponent will use similar calculations.

To approximate real-world deterrence one might look at boundary conditions for these equations. If one allows the utility of an attack followed by retaliation to approach $-\infty$ (utter destruction), the following occurs (if p is not equal to zero):

$$\lim_{u_1 \to -\infty} (p)(u_1) + (1-p)(u_2) = -\infty \qquad (1.2)$$

where u_1 is the utility of an attack followed by retaliation and u_2 is the utility of an attack not followed by retaliation.

In such a case, even if the probability of retaliation is tiny, as u_1 approaches negative infinity, expected value will approach negative infinity. If one can convince the aggressor that retaliation would be terribly harmful ($u_1 = -\infty$), demonstrating one's commitment to retaliation becomes less important. This suggests that the commitment problem is in some sense solved by making u_1 equal to negative infinity. One need only have enough credibility for the potential attacker to believe that *some* probability of retaliation exists. Any probability greater than zero fulfills this requirement.

While this may be mathematically elegant, experimental psychologists present considerable evidence suggesting that when the subjective utility derived from an event is extremely negative and the probability of occurrence near zero, most individuals act as if $p = 0$.[4]

---

[4] P. Slovic, B. Fischoff, and S. Lichten, "Behavioral Decision Theory," *Annual Review of Psychology,* Vol. 28, pp. 1-39.

For example, consider an individual contemplating a flight on a commercial airline from Los Angeles to New York. Assume that the individual considers his own death to have a utility equal to negative infinity. He recognizes that there is some, albeit low, probability that the aircraft will crash, killing all aboard. Assume that the reward for making the flight is not positive infinity. Expected value calculations (see equation 1.1) would produce an expected value of negative infinity. Clearly, he should choose not to fly. Of course, similar calculations produce an expected value of negative infinity for staying home also. One feels pity for this mortal rational actor who is faced with an expected value of negative infinity for all options, since as a mortal being, he always faces a probability of death greater than zero. This paradox is not central to this paper. For the purposes of this study one can conclude that individuals routinely engage in behavior which mathematical expected value calculations suggest is non-rational. One must conclude that below some threshold, small probabilities are treated as equal to zero.

Consequently, the deterrence theorist must convince the opponent that the utility from retaliation is equal to negative infinity *and* that the probability of retaliation is significantly above the discounting threshold. One cannot rely solely on making utility terribly negative. Since this threshold value cannot be established with confidence, the deterrer must work to ensure that the opponent believes the probability of retaliation is significantly above zero under *all* possible attack scenarios. Thus, the opponent's perception of one's commitment to retaliate remains a critical problem. If individuals did behave according to the logic of expected value, the requirements for successful deterrence would be quite minimal. Since they don't, planning a credible deterrent is a complicated and sophisticated endeavor. In policy terms, expected value calculations produce results which support the assured destruction proposition and the case for a small deterrent force with little flexibility.

The above analysis suggests that such support may be ill-founded. It supports those strategists who argue for greater flexibility in response options and some ability to limit damage from a nuclear war.

Greater flexibility suggests that regardless of how creative the potential aggressor is, he is unlikely to design an attack plan to which the deterrer cannot readily respond. Retaliation will follow any and all attacks.

The manipulation of probabilities also fails to produce the ultimate deterrent. For example, for a finite u_2:

$$\lim_{p \to 1} \quad (p)\ (u_1) + (1-p)\ (u_2) = (p)\ (u_1) \tag{1.3}$$

As p approaches 1, the expected value of an attack approaches the utility of retaliation (u_1). Thus, unless the attacker has a perverse utility function and derives some satisfaction from destruction of elements of his homeland, the expected value will become increasingly negative as p approaches 1. Conversely, no matter how large the utility from an attack followed by no retaliation (u_2) is, as 1-p approaches zero, this component of expected value also approaches zero. In the limiting case, $(1-p)\ (u_2) = 0$. Under these conditions any u_1 less than zero will deter, even retaliation with one ICBM. Clearly this is not plausible. A 100 percent probability that some unpleasant outcome will occur may be theoretically sufficient to deter. Empirically, it is not. Just as certain small probabilities are treated as equal to zero even when the associated utility is highly negative, perhaps barely negative utilities are treated equal to zero, independent of the associated probability.

Deterrence may be more a function of the difference between damage to one's own society and the aggressor's. This is precisely what is suggested when the argument is made that the Soviets would suffer significantly less in a nuclear war than the United States. Such a possibility undermines deterrence. Thus, the negative utility resulting from a successful U.S. retaliation directed against the Soviets is discounted by them according to the success of their initial attack:

$$u_1 = (u' - u'')$$

where

u' = Soviet casualties

and

u'' = American casualties

Assured destruction will fail as a deterrent if the Soviets have a greater cost tolerance. Even equal casualties would produce victory in Soviet eyes because they perceive themselves as being able to withstand greater hardship. Unfortunately, many Americans fall prey to political solipsism believing that all nations share American values. For example, according to American expected value calculations, the North Vietnamese were destined to lose the war because their casualties were several times those of American forces. Through their willingness to pay a much higher price, the Vietnamese persevered and won the war. Similarly, the Soviets might believe political victory through nuclear war is possible because of their greater cost tolerance. For deterrence to succeed, $u' - u''$ should be greater than zero. If one accepts the assured destruction assumption that $u' = u''$ = negative infinity, the difference between u' and u'' is indeterminate. Since subtracting negative infinity from negative infinity does not produce a discrete result, such levels of destruction might be totally disregarded. Alternatively, one might let $u' = u'' = 0$; then $(u' - u'') = (0 - 0) = 0$. Most nations, however, would prefer mutual destruction to their own in isolation. Death does not seem as bad when one knows that the enemy received a similar reward. Letting the utility of destruction of each society equal zero or negative infinity is unacceptable because in each case the death of both does not produce a higher utility than solitary death. Only assigning actual numerical values avoids these problems. Unfortunately, as discussed above, there is no basis for assigning such values.

The analysis becomes more complex when one simultaneously considers decisionmaking by two countries, A and B.

	A: WAIT	A: STRIKE (q)
B: WAIT	u_{11}, v_{11}	u_{12}, v_{12}
B: STRIKE (p)	u_{21}, v_{21}	u_{22}, v_{22}

where:

u = B's utilities in A's eyes
v = A's utilities in B's eyes
q = Probability of A attacking (in F
p = Probability of B attacking (in

Fig. 1.2 -- Decision Matrix for Countries A and B[5]

Assigning different variables allows one to consider deterrence between countries possessing divergent utility functions. A's choices are to wait or attack. Whether A waits or attacks there is, in his opinion, p probability that B will attack. If A perceives this as significantly high, he will "have no choice" but to attack. Conversely, there is 1-p probability that B will not attack and peace will continue. On the

[5]This figure and equations 1.4 through 1.9 are derived from Daniel Ellsberg, *The Crude Analysis of Strategic Choices,* P-2183, Santa Monica, California: The Rand Corporation, 1960.

other hand, if B waits he perceives q probability that A will attack and 1-q probability that no attack will occur.

A plausible preference ordering for A is the following, as estimated by B:

1. No war (v_{11})
2. Strike while B waits (v_{12})
3. Both strike (v_{22})
4. Wait while B strikes (v_{21})

Similarly for B,

1. No war (u_{11})
2. Strike while A waits (u_{21})
3. Both strike (u_{22})
4. Wait while A strikes (u_{12})

Assuming $p = q = .5$, deterrence succeeds when the combined utility of cases 1 and 4 is greater than the combined utility of cases 2 and 3.

A will not attack when:

$$(v_{11} + v_{21}) > (v_{12} + v_{22}) \qquad (1.4)$$

B will not attack when:

$$(u_{11} + u_{12}) > (u_{21} + u_{22}) \qquad (1.5)$$

More generally, the deterrer wishes to ensure that:

$$[(1\text{-}p)v_{11} + pv_{21}] > [(1\text{-}p)v_{12} + pv_{22}] \qquad (1.6)$$

(A chooses to wait)

AND

$$[(1\text{-}q)u_{11} + qu_{12}] > [(1\text{-}q)u_{21} + qu_{22}] \qquad (1.7)$$

(B chooses to wait)

Deterrence succeeds when *both* conditions are met. Each side must be convinced that the two conditions exist simultaneously. Each must be capable of solving inequalities 1.6 and 1.7, using both their own utility values/subjective probabilities *and* the opponent's. One cannot impose one's own utility function on the opponent. If so, egregious errors may result. First, each must believe that the other is not an easy prize (i.e., the opponent will retaliate if attacked). Second, each must believe that the other is not planning to attack; which is to say that each must believe that it is perceived as neither a wimp nor bully. This second requirement suggests that the U.S. move with care in acquiring first-strike capabilities or endorsing a strategy which terrifies the Soviets. For example, a strategy designed to destroy the Soviet government *per se*, might so unnerve Soviet leaders that rather than be deterred, they would choose to launch an attack. If this sequence of events occurred, one could say that the American deterrent strategy destroyed rather than enhanced deterrence. Both "wimps" and "bullies" attract aggression; wimps because they are cost-free targets and bullies because they provoke preemptive attacks. Deterrence succeeds best when each country perceives the other as tough but peaceful until assaulted.

Deterrence succeeds when the expected value from waiting minus the expected value from striking is greater than zero.

PARTY A

$$\underbrace{[(1-p)v_{11} + pv_{21}]}_{\text{(e.v. waiting)}} - \underbrace{[(1-p)v_{12} + pv_{22}]}_{\text{(e.v. attacking)}} > 0 \qquad (1.8)$$

and

PARTY B

$$\underbrace{[(1-q)u_{11} + qu_{12}]}_{\text{(e.v. waiting)}} - \underbrace{[(1-q)u_{21} + qu_{22}]}_{\text{(e.v. attacking)}} > 0 \qquad (1.9)$$

Ideally, these conditions will be met with a large margin for error. As stated above, deterrence is less and less likely to succeed as this difference approaches zero.

Deterrence theory is the study of mutual manipulation by threat. Each actor seeks to manipulate the behavior of the other by influencing his perceptions about the original actor's capabilities and intentions. Actor A will seek to massage (color) actor B's entire world view. Actor A very much wants to produce a convincing hallucination for Actor B, yielding "ideal" behavior. Furthermore, one actor will seek through declarations and precedent-setting examples to demonstrate that his behavior is contingent upon the other. Usually, both the promise of rewards and threat of punishment are held out. At the same time, Actor B is attempting a similar manipulation. One can, consequently, run into an infinite iteration problem. A thinks about B thinking about A's attempts to out-guess him, etc. Let it suffice to say that each country seeks to decrease the size of the opponent decisionmaker's domain through deterrent threats.

In Figure 1.3, the deterrer has succeeded in removing nuclear war from the leader's daily decision domain. No daily calculations on the expected value from nuclear war are needed. This is precisely what happens in Washington, D.C. and Moscow every day. Nuclear war is simply not discussed by national leaders on a daily basis.

A dramatic event or gradual shift in the ratio of forces can produce a discontinuity which brings nuclear war into the decision domain and onto the menu of options.

This discontinuity causes nuclear war to be included in the list of options considered by the leaders at t + 1. Such a discontinuity must precede any actual calculations. This is one reason why nations increase their alert levels during crises. They recognize the possibility that nuclear war has now been included on the list of daily options of the opponent and that, consequently, nuclear attack is more probable.

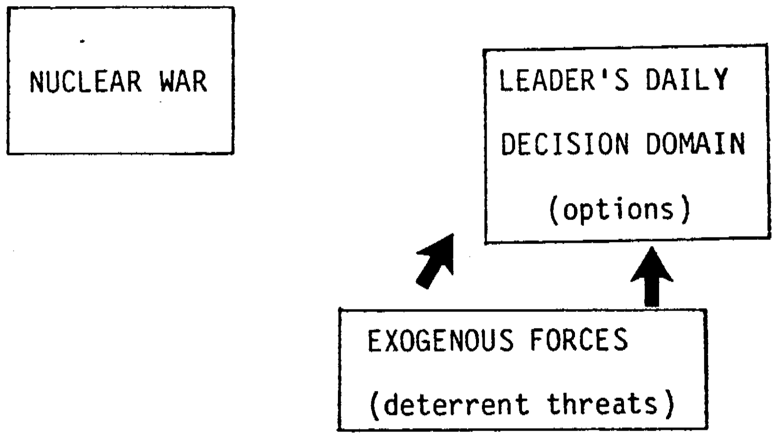

Fig. 1.3 -- Leader's Peacetime Decision Domain

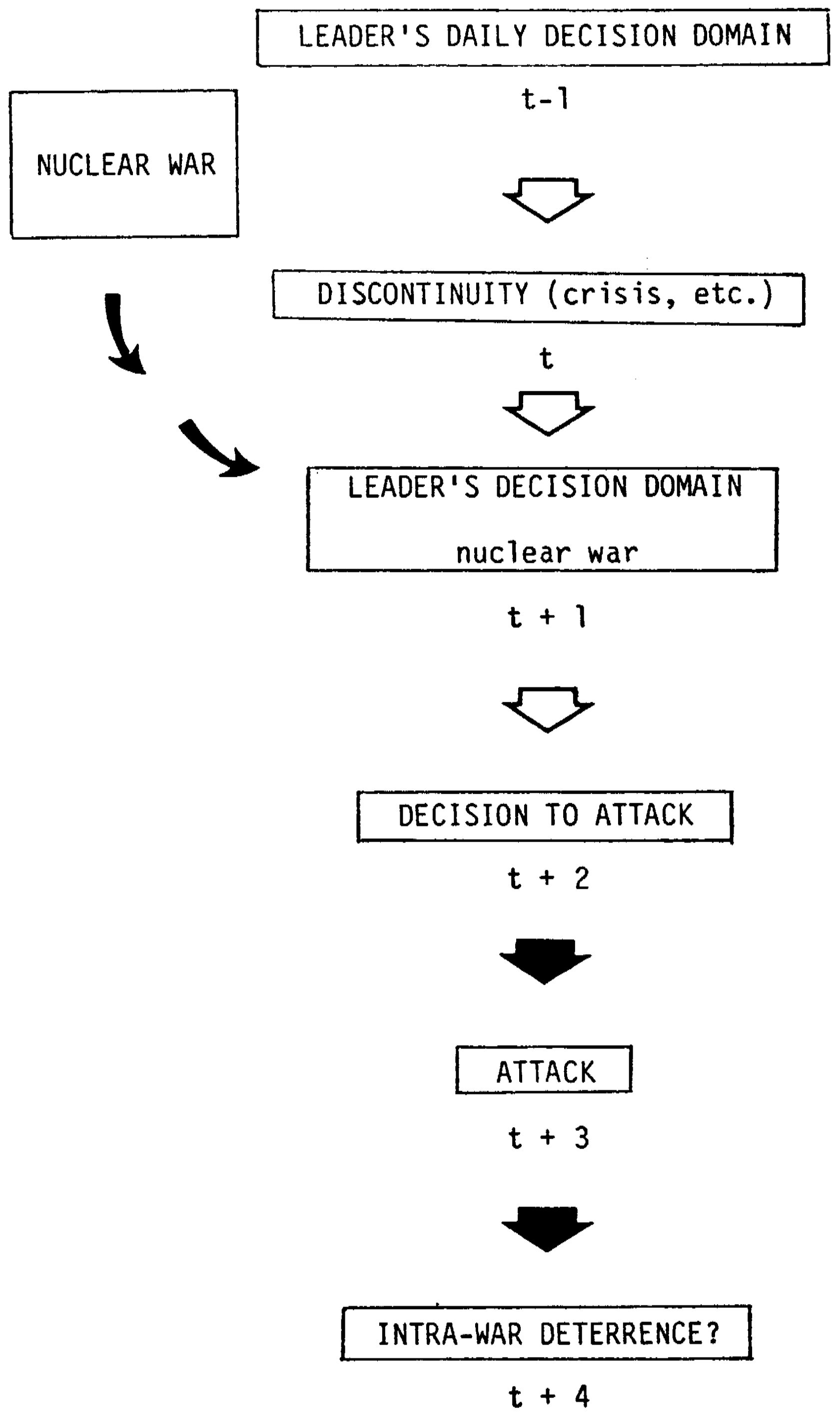

Fig. 1.4 -- Leader's Wartime Decision Domain

UNCERTAINTY AND THE PROBLEM OF COMMITMENT

The ability of one actor to achieve his goals is in large part a function of the preferences and decisions of the other actor. This occurs under conditions which are characterized by neither pure cooperation nor pure conflict but rather by a combination of the two.[6]

Consider the Soviet-American relationship. Although these countries have conflicting ideologies and interests on many dimensions, they do have an overriding mutual interest. Since each country is capable of doing great harm to the other with nuclear weapons, it is in their mutual interest to manage their differences with great care and to avoid confrontation.

Deterrence is threatened when significant doubts exist about one state's commitment to fulfill its threats. Doubt about the ability and/or willingness of an adversary to retaliate encourages aggression because it suggests the possibility of political gains achieved through low-cost or cost-free nuclear attack.

For example, American leaders seek to convince the Soviets that the United States will not attack the Soviet Union without provocation. Using inequality (1.6) one can see that if the U.S. is country B and the Soviet Union is country A, then the United States seeks to convince the Soviets that the following inequality expresses U.S. expected values.

$$[(1-q)u_{11} + qu_{12}] > [(1-q)u_{21} + qu_{22}] \qquad (1.7)$$

(e.v. from waiting) > (e.v. from attacking)

Since the expected value from waiting is greater than the expected value from attacking for the United States, it will not attack without provocation. The Soviet leaders, however, must have *no doubt* that in the event of a nuclear attack against the United States, retaliation will be swift and severe. The Soviets must believe that the expected

[6] Thomas C. Schelling, *The Strategy of Conflict*, Cambridge, Mass: Harvard University Press, 1960.

value from retaliation is greater than the expected value from surrender for the United States *under all conditions.*

One approach attempts to convince the Soviet leader that his decision to attack would set in motion a train of events which is independent of the commitment or competence of any one individual. Retaliation is presented as an automatic and institutionalized force which has its own dynamics, its own laws. The desire to make retaliation automatic is behind a suggestion, for example, to give military leaders the authority to order nuclear retaliation. Military men are perceived as more likely than civilian leaders to order retaliation. A potential aggressor would be more impressed by military control (probability of retaliation, p, increases) and, therefore, less likely to attack. The doomsday machine discussed by Schelling and illustrated in *Doctor Strangelove* is based on a similar logic. In the case of a doomsday machine, removal of the decision to retaliate from human hands improves deterrence. Once the machine detects an attack it launches a retaliation (Schelling's machine) or detonates enormous "dirty" warheads buried around the country which produce a radioactive cloud so deadly that all life on the Earth is killed (the machine in *Doctor Strangelove).* Both are launched without human intervention. Indeed, any human attempt to interfere or prevent retaliation will cause the machine to detonate its weapons. Computer error or the accidental detonation or launch of a few weapons would also cause global destruction. While such a machine is ideal for solving the problem of commitment, it makes accidental nuclear war more likely.

A less risky solution is to present retaliation as something not wholly under the control of the national leader.[7] The national leader might say to his opponents, "Look, even if I want to surrender, I may not be able to." The decentralization of command and control over nuclear weapons which makes a preemptive strike unlikely to succeed is the same decentralization which would make it possible for military commanders to order "unauthorized" retaliation. To some extent, this "lack of control" strengthens deterrence.

---

[7]See Thomas C. Schelling's "Surprise Attack and Disarmament" in Klauss Knorr, editor, *NATO and American Security,* Princeton: Princeton University Press, 1959 and his *Arms and Influence,* New Haven: Yale University Press, 1966.

Consider Schelling's "climbers on the mountain" analogy. In the classic formulation of nuclear deterrence, the nations were pictured as two individuals who stand roped together at a firm, demarcated abyss. The deterrer warns that in the event of a limited nuclear attack, he will respond by throwing himself off the cliff (launch an unlimited retaliation), dragging the other party with him and both to their deaths. Schelling suggests that instead one should think of the brink of war as akin to a mountain whose slope becomes steeper, less predictable, and more treacherous as one proceeds down. There is no clear point where disaster beckons. The danger of falling does not increase in a consistent or predictable way as one climbs down. At some locations the footing is good and rock solid; yet, without warning the rock can be brittle, loose and unstable. Thus, all that one knows with certainty is that the probability of disaster increases in some incalculable way as one moves down. Under this uncertainty-based deterrent, the threat is not "I will jump, dragging you with me," but "I will take one more step, increasing the probability that I will fall and despite my frantic efforts to save myself, I will continue falling, dragging both of us to the bottom and certain death."

One might view this as the opposite of Herman Kahn's[8] escalatory ladder. Instead of climbing a ladder of escalation, one descends it. Each rung represents a more severe escalation than the one above it. The bottom of the ladder and ultimate stage of escalation may be reached gracefully one rung at a time or by taking a tumble off the ladder. It really makes no difference how one reaches the bottom (global nuclear war). In either case, the ultimate stage is reached and both sides lose. The mountain analogy suggests that while one may respond to the aggressor's initial limited attack in a limited fashion, eventually an unseen and unknowable threshold is reached. Once this threshold is crossed, control is lost along with one's society.

Schelling's analogy is less compelling to those who believe that mutual suicide is not a particularly believable outcome for nuclear war. The image of two actors roped together implicitly endorses the assured destruction proposition. Yet, the link may not be this strong. Nuclear

[8]Herman Kahn, *On Escalation,* New York: Praeger, 1961.

threats may be secondary to actual escalatory options. Thus, an alternate approach would have the two climbers proceed down the mountain independently and unroped. The leader challenges the other to follow him down. Eventually, a point is reached where the less skilled climber slips and falls to his death, *alone.* The superior climber survives and *wins,* just as the country with superior war-fighting and escalatory potential might prevail, achieving escalatory dominance. At the least, this produces war termination under favorable conditions. At most, this produces total economic, political, and military defeat of the adversary.

Schelling's analogy suggests to assured destruction proponents that the rationality requirements of nuclear deterrence are quite innocuous. One need only be able to recognize the extreme danger of milling about on mountains. The statesman need not calculate where the discrete threshold for disaster is. Indeed, he could not even if he wished to. He only knows that the gray areas of confrontation and nuclear war are filled with uncertainty, uncontrollable forces, and hidden dangers. The prudent statesman avoids such uncertainties. Alas, even prudent statesmen grow accustomed to living on the dangerous mountain. Danger and uncertainty may not be enough to deter statesmen who live under the nuclear threat on a daily basis. Rather, the warfighters' flexible targeting options and other war-waging abilities may be necessary to convince the opponent that one is the better "climber."

The resistance of the uncertainty "problem" to solution does, however, enhance deterrence. If the statesman could place every variable of nuclear war into a strategic calculus and great confidence was held in the equation, a national leader would be more likely to see nuclear war as a policy option than in the case where significant uncertainties exist about weapon reliability and accuracy, command system reliability, personnel competence, and the dynamics of war escalation.

Just as the climber cannot predict the consequences of his steps down the mountain, the statesman cannot be confident of the way that diplomatic posturing, hot-line threats or actual nuclear exchanges will affect the adversary's behavior. As long as the statesman recognizes that war is a phenomenon with its own dynamics and that it may produce actions which cannot logically be linked to the rationale for the war, the credibility of deterrence is enhanced through uncertainty.

Uncertainty helps solve the commitment problem. The aggressor cannot confidently call the deterrer's bluff if the deterrer is in less than total control of his own forces. Even if a nation's leaders are willing to lose face and have their bluff called (e.g., surrender after the destruction of ICBM fields or other targets), retaliation will probably still occur. The aggressor cannot count on the victim's command and communication systems surviving even a limited attack. If communications are disrupted, authority for retaliation will cascade downward to field commanders. Retaliation will probably occur before the political leadership reestablishes control and orders the cessation of hostilities. The deterrent, therefore, is somewhat independent of the national leadership's preferences, personal characteristics (dovish vs. hawkish) or competence.

Another issue is that of technical uncertainty. Can the nuclear forces in each country actually ride out a surprise attack? Are the weapons numerous and dispersed? Are the commanders and launch crews protected from the effects of nuclear weapons? Is a high alert level maintained? Will a significant number of warheads reach their targets after an attack? Will a single technological breakthrough render all legs of the triad (bombers, submarines, and land-based ICBMs) obsolete overnight--the strategic breakout problem?

Retaliatory uncertainty will affect the attacker's calculus. The more the potential attacker questions the victim's ability to retaliate, the more likely an attack is. For example, an attacker is likely to assign probabilities to many types of retaliation.

Thus, the attacker's expected value calculations will be based on an analysis of the severity and likelihood of different types of retaliation.

The attacker must also recognize that few of his weapons are likely to perform as well in combat as under perfect test conditions. Neither the United States nor the Soviet Union has ever launched more than a handful of missiles simultaneously. It is unlikely that the attacker could launch hundreds of missiles within seconds of one another with an untried C^3I system. Will the personnel respond in a timely and

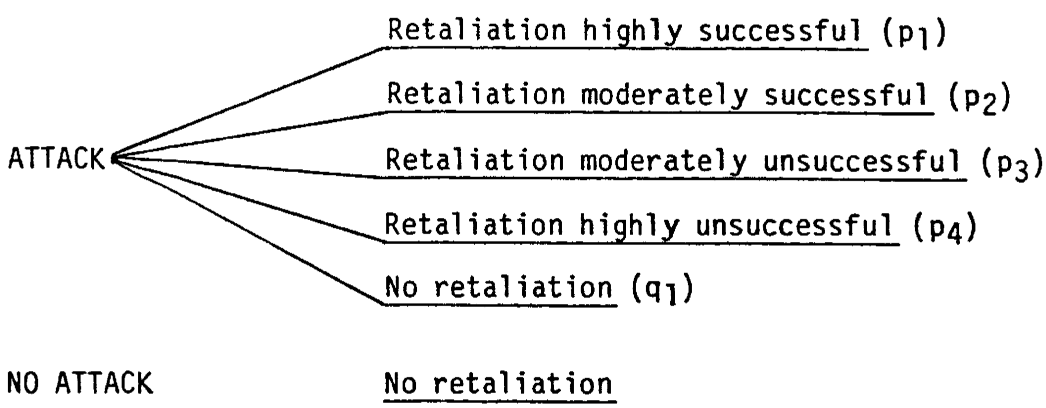

Fig. 1.5 -- Attacker's Decision Tree Under Conditions of Uncertainty

competent manner? What if many of the launch crews insist on time-consuming verification of the launch orders or balk at the idea of launching a real nuclear war? What if the victim responds by launching on warning? Even worse, what if the victim responds to strategic or tactical warning signals and launches a preemptive attack? Less dramatically, the victim may respond by flushing his bombers to hundreds of airfields, sending submarines out to sea and placing all forces on an alert status. Any attack under these conditions is likely to result in a launch on warning, negating the advantage of striking first.

For the victim, his weapons are not the only targets which must have the ability to survive the initial stages of a nuclear war. His command and control systems must be equally survivable. For example, a very limited attack against key command centers and the known patrol areas of airborne command aircraft might achieve total surprise. Would such an attack significantly weaken the retaliation through the loss of control and inability to transmit the attack order?

A more exotic and highly controversial attack would rely, not on the blast and heat effects of nuclear weapons detonated on the ground or at a few thousand feet above the ground, but on the electromagnetic pulse (EMP) effects from large yield weapons detonated in orbit, a few hundred kilometers in space. In theory, ten explosions would produce an electromagnetic surge exceeding 50,000 volts per meter over the entire United States. This surge would overload integrated circuits, disabling all unprotected electrical components. Military antennae, communication wires, aluminum aircraft bodies, and many other elements of the U.S. C^3I system are ideal receptors for EMP. The uncertainties associated with an EMP attack, however, are even greater than those associated with the more common counterforce attack. It is entirely possible that an EMP attack would do little more than provide an awesome fireworks display and warning of dishonorable intentions on the part of the opponent. Such an "attack" would be an act of war of the most serious kind, likely to provoke nuclear retaliation.

Mutual concern over command and control vulnerability, like ICBM vulnerability, can undermine crisis stability. During a severe crisis each country might feel compelled to launch first rather than risk a *de facto* loss of its entire nuclear force. Once a nuclear war had begun, the Soviets might fear that the United States would use its surviving C^3I assets to track and destroy the surviving elements of the Soviet C^3I system. Each side would feel compelled to destroy the other's command assets, lest its be destroyed. Thus, the same arguments for survivable retaliatory weapons apply here. Survivable command systems and weapons are both essential pillars of deterrence.

The preceding discussion implies that avoidance of nuclear war is the overriding concern of national leaders and that they act in a calculated manner, always seeking to minimize the danger of nuclear conflict. Actual crisis behavior suggests that this is rarely the case. National leaders have many, often conflicting, responsibilities and the avoidance of war is not necessarily the most compelling imperative. Thus, the use of deterrence theory must be grounded in a more traditional appreciation of the historic tension between war avoidance and the protection of national interests.

RAND/P-6916 SOME THOUGHTS ON DETERRENCE Alan Vick