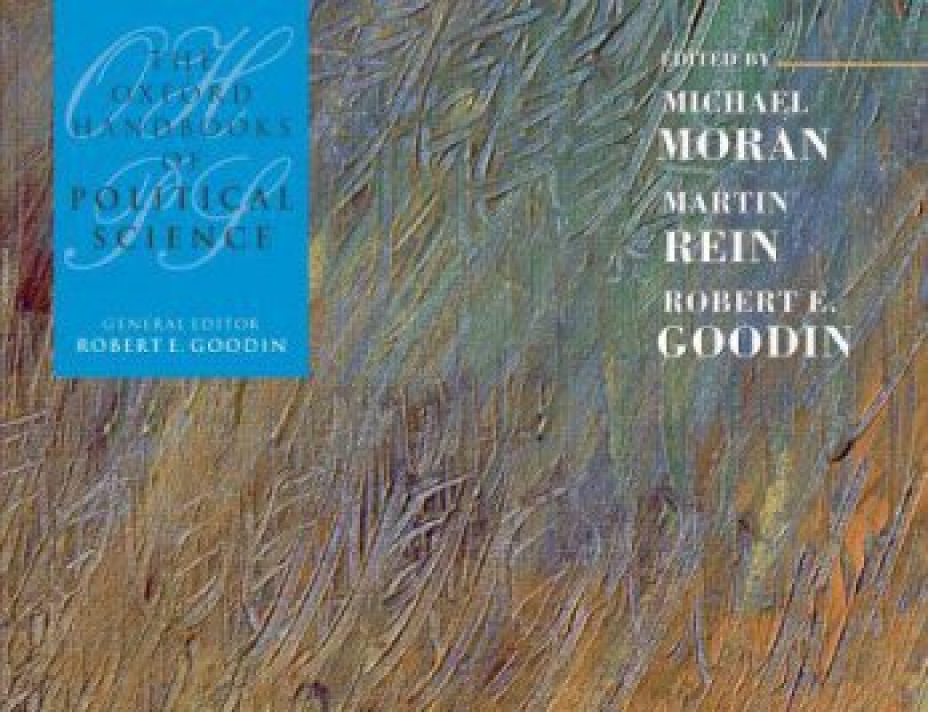

THE OXFORD HANDBOOKS OF POLITICAL SCIENCE

GENERAL EDITOR ROBERT E. GOODIN

EDITED BY

MICHAEL MORAN

MARTIN REIN

ROBERT E. GOODIN

The Oxford Handbook *of* PUBLIC POLICY

number of artificially simplified tests. Recall that CEP, a distance measured in nautical miles or feet, is the radius of a circle around the target where 50 per cent of the warheads are expected to fall if a large number of test firings were conducted. Some 50 per cent would likely fall outside this radius.[22] Accuracy depends on the gravitational and electromagnetic field of a missile flight path, precise calibration of the inertial guidance system of a weapon, that the re-entry vehicle does not get thrown off course by debris when it re-enters the atmosphere, and so on. The tests that were used to estimate US missile accuracy were conducted on east to west flight paths, over what is known as the Western Test Range, while a US ballistic missile flight against the USSR during the cold war would have gone over the North Pole and over longer ranges—these missiles would experience different gravitational and electromagnetic forces. Moreover, the missiles that are used in these flight tests are specially prepared and "modified" for the tests, so that they are in better working condition than the missiles that actually sit in silos or on submarines (MacKenzie 1990, 344).[23] The missile warhead lands in the test area and the number that is eventually given for CEP of a particular missile type depends on a statistical analysis of a number of these tests. To take uncertainty into account, there are "safety factor" formulas that are apparently used by systems analysts for CEP (MacKenzie 1990, 419). Yet the CEP number is generally taken as a given when inputted into systems analysis calculations.

Ironically, uncertainty, and the sources of uncertainty with respect to CEP were sometimes discussed in great detail by policy modelers and then ignored. For example, the Congressional Budget Office (CBO) produced a number of widely used papers examining US strategic nuclear forces in the 1970s and 1980s. The CBO report was careful to make the problems and uncertainty with the data explicit and also to note that even if more tests were conducted in order to increase confidence in the CEP figures used in the analysis, "actual" nuclear war would be quite different from the tests:

> A very significant consideration for attack planning is the great uncertainty surrounding the actual accuracy of any given guidance technology. This uncertainty results in part from the limited number of tests a missile system undergoes to verify its accuracy potential. Gaining high confidence in estimates of a missile CEP would require a large number of tests for each missile and for each change in its guidance system. Such testing is constrained, however, by the limited resources that can be devoted to the very expensive task of missile testing. Moreover, actual operational performance can be degraded by variable atmospheric conditions and small perturbations in the earth's gravitational field. As a result, actual CEPs can only be estimated within a fairly large range of uncertainty, and any assessment of the damage that an

[22] Lynn Eden suggested to me that this is an odd locution: it is *circular* error probable although weapons would not fall in a circle but in more of an elliptical pattern.

[23] One could respond that because of these areas of uncertainty, one needs to do more tests. In fact, those who do not want to halt nuclear tests or tests of delivery vehicles and components argue that periodic testing of nuclear weapons and delivery vehicles is necessary to ensure that the weapons will be reliable and that the assumptions about performance are accurate. Yet, even if testing advocates had their way, tests would still be stylized simply because to get the necessary measurements, tests must be conducted under "artificial" and stylized conditions.

attack can be expected to cause must take into account the uncertainties surrounding these operational accuracies. (CBO 1978*a*, 10 11)

Yet, although data reported by the CBO as the basis for their calculations were frequently used by other analysts, the explicit cautions expressed in the CBO reports, including the one quoted above, are rarely reproduced. Thus, the problem of uncertain inputs being used as hard numbers was exacerbated by the tendency of analysts to simply repeat earlier estimates made or given by respected sources (Crawford 1987). Uncertainty was thus acknowledged and then forgotten or erased and turned into hard and certain numbers which became the basis for other calculations. Simulations were taken to be real and accurate, when they were highly constructed and likely to be far from accurate; the analysts knew this and proceeded anyway.

Further, uncertainty was magnified and masked when classified and public estimates frequently based on projections of *future* capabilities of the USSR rather than on what was known or presumed to be the current capability. There were enormous questions about contemporary Soviet military capabilities; those uncertainties were even greater if Soviet capabilities were projected into the future. For example, projections of future Soviet capabilities that never actually materialized were the basis of the highly publicized bomber and missile gaps. Classified estimates in NIEs and operations research studies also, as a general rule, proceeded on the numbers projecting future capabilities. For instance, the 1964 classified study of damage limitation estimated US and Soviet capabilities for 1970 (DDR&E 1964*a*) but no one could know for sure what the Soviet arsenal would look like in six years and the basis for such projections was often never specified. Even the use of the term "projection" in the estimates connotes a systematic and empirically based number when what was given was often simply a guess of what the Soviets might be capable of doing in the future.

Omission and elision. "It is a serious pitfall," Quade (1968*b*, 359) argues, "for the analyst to concentrate so completely on the purely objective and scientific aspects of his analysis that he neglects the substantive elements or fails to handle them with understanding." Despite this caution, issues and numbers that are important for understanding the capabilities and effects of nuclear weapons are often omitted during the process of systems analysis. Four examples—the persistence in ignoring or downplaying the thermal effects of nuclear weapons, the omission of command and control in many models, the problem of fratricide, and the lack of reference to human bodies—illustrate the sort of omissions that characterize nuclear modeling.

As Lynn Eden (2004) shows in her masterful account, nuclear planners focused on blast effects, despite the fact that the thermal effects of nuclear weapons are enormous: when combined with the wind that nuclear explosions generate, huge fires would be expected in cities. As Eden demonstrates, blast effects are certainly important, but when trying to model the destruction of nuclear missile silos or other hard structures and when weapons planners talk about targeting cities and industrial targets, they usually took *only* blast effects into account. For example, the RAND Corporation SNAPPER Nuclear Damage Assessment Model focuses on blast effects. SNAPPER was used by the CBO (1978*b*) for its modeling, although the CBO noted that "*Secondary* effects from a nuclear blast, such as fire or shorts in electrical systems, can damage machinery just as effectively as primary effects can" (emphasis added; CBO 1978*b*, 47). Yet, depending on the dominant building materials and other conditions, the area of damage from a nuclear blast in a city and perhaps even against weapons will likely be much smaller than the area damaged by heat and firestorms. Firestorms did significant damage in Hiroshima and Nagasaki and there was other evidence that thermal effects of nuclear weapons would be significant. Still, nuclear modelers preferred to focus on blast because they believed blast effects were easier to predict and model. This example of a preference to model blast effects is taken from a now declassified memorandum to President Kennedy, and occurs in a discussion of the kill distance of anti-ballistic missiles against a swarm of incoming nuclear warheads and decoys by RAND Corporation experts Edward Teller and John Foster:

> Suppose the kill distance of the defensive warhead could be vastly increased made comparable to the size of the swarm. The decoys could become ineffectual.
>
> If there were multiple warheads they could all be killed in one blow....
>
> How can the kill distance of a nuclear warhead be made so large? Is such a warhead development possible? The answer is that it may not be necessary to do anything to the warhead. The kill distance with present warheads might be big enough *and we just don't know it.* It is an important fact that the science of the effects of nuclear explosions on targets is in a much more rudimentary state than the science of nuclear weapons themselves. Because we know so little about effects and because we do not know the detailed construction of the Soviet ICBM, we are forced to base our estimate of the kill distance on the most direct, best understood, and therefore most reliable effects of the explosion. It is this way [deletion]....
>
> Aren't the Soviets, like us, forced to be conservative in their AICBM [anti ICBM] planning? (Teller and Foster 1961, 3, 5)

Thus as Teller and Foster imply, the fact that a firestorm would likely destroy a vast area was not taken into account because analysts were focused on the blast effects of nuclear weapons, and the result of considering other effects "secondary" is that more nuclear weapons would be targeted on an area such as a city, to produce damage to a certain level of blast. The idea that one needs more weapons often leads to building them, and then the other side may build weapons to be able to target those weapons, and so on.

Command and control of nuclear forces was also often omitted from analysis by the assumption that it would work flawlessly or at least quite well. There were about thirty-six nuclear command posts in the USA and fifty in the USSR in the mid-1980s (Arkin and Fieldhouse 1985, 93; Blair 1985). As Ball (1986*a*, 19) argues, "Escalation Control requires U.S. strategic nuclear forces be supported by a survivable C^3I system with sufficient endurance to maintain control through some extended period of protracted conflict." But as Ball shows, US C^3I is "subject to certain critical vulnerabilities" which call into question the ability to follow through with war fighting scenarios. Despite command and communication redundancies and other precautions, an attack on all of these command posts would likely hinder political leaders' ability to launch nuclear weapons, assess damage to their own and the other side, or terminate a nuclear war once it was begun. The smooth and effective functioning of C^3I, essential for all nuclear war scenarios, is assumed in most systems analysis of second strike retaliation, despite the fact that C^3I is quite vulnerable to disruption.[24]

Analysts also sometimes acknowledged and then proceeded to omit from their calculations the possibility of fratricide—that the detonation of one of your weapons could disable another of your weapons—from their analysis. Specifically, to increase the overall probability of kill (OPK) against a target, nuclear weapons planners often allocate more than one nuclear weapon to it. "To hedge against massive failures of an entire weapon type, weapons would be cross targeted by different delivery systems" (Postol 1987, 380). Cross-targeting raises the possibility of fratricide because the first weapon to explode will create a fireball and dust cloud. "If the second cross-targeted booster did not fail in flight to the target, its warheads would arrive next, perhaps minutes or fractions of minutes after the arrival of the first. . . . Some of the warheads might be damaged or destroyed if they encountered the debris clouds from the earlier detonations, but from the point of view of the targeter that might be unimportant, because the warheads would be cross-targeted mainly to make it highly probable that the targets of interest were struck" (Postol 1987, 389). But according to the CBO, "It is possible that no more than one warhead could be successfully detonated over each target. Other nuclear effects, such as intense heat and dust clouds, could be lethal to subsequent warheads even if first round weapons were burst above the surface to avoid the throwing of ground debris into the air" (CBO 1978*a*, 12). Moreover, "Uncertainties about fratricide will probably never be settled. For one thing the prohibition on atmospheric testing prevents real world evaluation of a modern warhead's ability to withstand the various effects of a nuclear explosion" (CBO 1978*a*, 13). Despite these significant concerns, fratricide is often left out, or minimized in calculations by strategists. The result is that the "models" are less and less removed from the "reality" of the weapons effects, even as the conclusions of models based on this optimistic assumption create yet another sort of reality.

Finally, as Cohn (1987) and Gusterson (1996) have noted, one of the most glaring omissions is the frequent lack of clear references to what nuclear weapons do to humans. Of course one of the main points of using nuclear weapons is to kill people. Calculations about "countervalue" strikes against population centers do discuss the casualties associated with nuclear weapons use (e.g. OTA 1979; Batcher 2004). But, apart from the early research on the effectiveness of civil defense, many of the counterforce calculations proceed as if there were no human injuries or deaths from counterforce nuclear exchanges. Indeed, the intentional and inadvertent

[24] Though command bunkers and other elements of C3I are "hardened" against blast, transient electronic effects (TREE), and electromagnetic pulse (EMP), they are still vulnerable to direct hits.

release of sometimes high doses of radioactive substances during nuclear weapons tests and as part of the program of human radiation experiments undertaken in the USA during the first fifteen years of the cold war (see Hilts 1994, 1995, 1996; Wald 1997) could lead to speculation that human life itself was discounted by some planners.

Arbitrariness. The inputs to policy modeling should be based on non-arbitrary considerations. Yet, modeling inputs used as baselines in nuclear systems analysis, and that seem relatively uncontroversial, such as the size of the ICBM arsenal, for CEP, and the criteria of second strike survivability, were all too often arbitrary. An initial arbitrary assumption may appear uncontroversial, but the effects of the initial policy choice ripple through subsequent analysis.

For example, there was no compelling military or scientific reason why the US ICBM arsenal was set at 1,000 missiles (Ball 1980, 209–10). In 1974 nuclear scientist Herbert York asked Alfred Rockefeller, chief of the Presentations Division of the Space and Missile Systems Organization of the air force, to explain how the size of the US ICBM force was determined to be 1,000 in the mid-1950s, suggesting that its number was essentially "a natural one, and not decided by anybody consciously" (York 1974). Rockefeller replied to York, "I agree with you on the interpretation of the number 1000. Basically, it is a nice round number which would be equally applicable to an aircraft procurement.... the number 1000 was a natural one. A nice base figure to calculate cost on" (Rockefeller 1974).

Similarly, the criterion used by NATO countries for accuracy CEP is 50 per cent probability of the warhead landing within a radius expressed in nautical miles or feet. According to this criterion, 50 per cent of the warheads land somewhere outside that radius. Again, this distance is calculated based on several test firings of the weapon, and the classified results of tests include confidence intervals and an error budget of the causes of inaccuracy (Mackenzie 1990, 348–9).[25] So, although the number for CEP is *expressed* as a distance, the circular error probable figure is a *probability* for landing within a certain distance. Yet, the choice of 50 per cent is essentially arbitrary. Why does NATO use 50 per cent as the probability? Clearly, if a different criterion were used, the distance would be different, altering one's perception of the missile accuracy, and therefore, likely altering the number of weapons procured. Why not use a different criterion, for instance 80 or 90 per cent, which would be more consistent with the numbers for reliability of missiles and warheads? Weapons would appear to be less accurate if CEP were 80 per cent and more accurate if it were 21 per cent, the figure the Soviets used for CEP.[26]

Other figures, taken for granted at the time as not arbitrary but as "reasonable and essential," were the criteria used to assess when deterrence would be accomplished. McNamara's Department of Defense asserted that deterrence would be accomplished

[25] Mackenzie (1990, 367 8) notes how CEP confidence intervals were viewed differently and CEP numbers adjusted when the air force wanted to make their nuclear weapons appear more accurate than navy weapons.

[26] Because the Soviet criterion for CEP was a 21% probability for landing within the radius they could expect 79% of their weapons to land outside that radius.

with the guarantee of 400 (later revised to 200) equivalent megatons for a second strike—that is, the USA should be able to inflict that amount of nuclear damage even after absorbing a first strike by the Soviet Union. McNamara told the Congress in 1965 that "it seems reasonable to assume the destruction of, say, one-quarter to one-third of its population and about two-thirds of its industrial capacity...would certainly represent intolerable punishment to any industrialized nation and thus should serve as an effective deterrent" (quoted in Ball 1986*b*, 69). Plans were developed to accomplish this level of destruction, and it was shown through systems analysis techniques that 400 EMT would do the job of visiting this much destruction on the Soviets. Yet, the number used by McNamara's Pentagon for "unacceptable" damage was essentially pulled out of the air and then the number of equivalent megatons necessary to do the job was calculated by looking at the diminishing marginal returns of doing more damage (Kaplan 1983, 316–18). These criteria were later changed. The *Annual Report of the Secretary of Defense, Fiscal Year 1969* (1969, 50) estimated that 400 EMT would be sufficient to destroy half of Soviet industry. The NUWEP-1 (Nuclear Weapons Employment Policy) of the USA in 1974 required nuclear weapons to destroy 70 per cent of the Soviet economic and industrial base needed to achieve economic recovery (Ball 1986*b*, 74). In 1978 US Secretary of Defense Harold Brown told Congress that it was "essential that we retain the capability at all times to inflict an unacceptable level of damage on the Soviet Union, including destruction of a minimum of 200 major Soviet cities" (quoted in Ball 1986*a*, 27). The CBO suggests that "Destruction of 80 percent of the industrial target set [of their 1,400 industrial target base] appears to be a reasonable objective" (1978*b*, 52).

Where did these numbers, which changed from administration to administration, come from? Why these numbers and not others? The requirements, and the arsenal built to accomplish them, appear to be arbitrary. No one knew for sure—or even with confidence—what would deter the decision makers of the Soviet Union or any other leadership. Maybe more, maybe significantly less destruction would be required. Arbitrariness and uncertainty is then glossed over by the use of words like "requirement," "reasonable," and "essential."

Sometimes opaqueness and arbitrariness were combined. For example, in a discussion of allocating weapons to targets, the CBO gave an example designed to illustrate damage expectancy: "[I]f the first target has a value of 1000 and the weapon Pk is 0.80, then, assuming 100 percent reliability, one bomb would destroy 800 units of target value. Allocating a second weapon to this target would result in additional value destroyed of 160 units. Therefore, this second weapon should be allocated to the first target before a target valued at 159 units is attacked" (CBO, 1978*b*, 53). But what is the unit of target value? A 1,000 what? How *should* targets be valued?

Such precise yet arbitrary inputs have the effect of making the activities of nuclear planners and preparations for nuclear war seem more accurate, but the consequences of the analysis were probably just the opposite. Even as nuclear analysts acknowledged uncertainty, and then developed and refined techniques for identifying and eliminating uncertainties from their models, they minimized the uncertainties that they did count and did not take into account other very important areas of uncertainty. In some equations, the mathematical "precision" of the models was accomplished by inserting numbers with little or no precise, "real" basis.

Implausible/"unrealistic" scenarios. Many of the systems analysis scenarios work on paper, but because they leave important effects out, or factor in unlikely events the scenarios are implausible. Three examples—issues of human reliability, the possibility of conducting nuclear and conventional war in an integrated and controlled manner, and the idea of "reprogramming" during nuclear war—are illustrative.

The question of human reliability was rarely discussed, much less factored into systems analysis. For example, there is one missile launch control center, operated by air force officers, for every 10 Minuteman and MX missiles (Blair 1985, 87). Thus, for MX missiles, which each have ten independently targeted warheads, one control center is responsible for launching 100 nuclear weapons. Commanders and systems analysts generally assume that humans will perform in a nuclear war environment as they were trained to function. Yet this is unknown and thus huge potential failures of reliability—humans may become ill or simply refuse to perform their duties—are rarely, if ever considered by systems analysts (see Dougherty 1987, 413–15). Omitting the discussion of human reliability has the effect of making the unrealistic assumption that human reliability will be perfect. Similarly, by not discussing Soviet reliability, one unrealistically assumes perfect reliability on their part.

US war planners also assumed it was possible to control escalation in nuclear war and developed plans for flexible, limited, and theater (local) nuclear war throughout the 1960s, 1970s, and 1980s. For example, Harold Brown presented Presidential Directive 59 in 1980 as a plan that would "integrate" strategic, theater, and tactical nuclear weapons use. "Our planning must provide a continuum of options, ranging from use of small numbers of strategic and/or theater nuclear weapons aimed at narrowly defined targets, to employment of large portions of our nuclear forces against a broad spectrum of targets" (Secretary of Defense 1980, 55). War plans included "integrating" nuclear, chemical, and biological weapons on the battlefield and discussed "selective employment of nuclear weapons against armored thrusts" (Joint Chiefs of Staff 1977, 85). US Army Field Manual 100–50 of March 1980, "Operations for Nuclear Capable Units," talked about training to "disperse" and "issue" tactical nuclear weapons rounds in a combat situation. "The U.S. has reviewed force levels and system requirements in an effort to achieve a TNF [Theater Nuclear Force] posture that will correct existing imbalances and provide credible, flexible responses, particularly at lower levels of nuclear warfare. Such a posture will provide timely accurate nuclear options for reinforcing deterrence outside the NATO area" (Joint Chiefs of Staff 1982, 29–30). The dubious assumption was clearly that the Soviet Union had conventional "superiority" which could be "corrected" by using tactical nuclear weapons. Yet, little attention was paid to the fact that the "employment" of tactical nuclear weapons on the "battlefield" could cause "tactical" nuclear war to escalate to all-out nuclear war.

Planning for "reprogramming" on the fly during nuclear war was also unrealistic. The idea was that nuclear weapons held in reserve would be retargeted to make up for weapons that failed to detonate or to retarget the targets not destroyed by the first round of weapons. Reprogramming is designed to increase the efficiency of nuclear targeting and boost damage expectancy, or the probability that the target will be hit and destroyed by a nuclear weapon. Reprogramming is often considered by nuclear planners who for instance, will make every effort to decrease the probability of fratricide by taking into consideration the height of burst and timing of follow-on nuclear bursts. However desirable it might be to increase efficiency, the scenario is implausible specifically because it assumes functioning damage assessment and command and control in a nuclear environment. On the other hand, the inadequacy of US C^3I in such a scenario was highlighted in Presidential Directive 59, where developing the requirements of counterforce were linked to making improvements in command, control, and communications (Ball 1986*b*, 78).

If the notions that humans were perfectly reliable, that nuclear weapons use could be limited to a battlefield, and that weapons could be reprogrammed in the midst of nuclear war are optimistic, there was also a tendency to emphasize worst-case scenarios—that the other side will do better and your own force worse in a nuclear war. This is known as being conservative or hedging. The tendency to think in terms of worst-case scenarios was reinforced by inferring an adversary's intentions from their military capabilities.[27] And oddly enough, a worst-case bias and hedging often occurs alongside a tendency to assume that things will go according to plan (that your equipment will function according to plan). For example, as mentioned earlier, it is common to omit command, control, and communications from nuclear systems analysis efforts at modeling nuclear war, perhaps because most assume perfect C^3I (see e.g. Salman, Sullivan, and Van Evera 1989, 191). "Conservative military planners tend to base their calculations on factors that can be either controlled or predicted, and to make pessimistic assumptions where control or prediction are impossible" (OTA 1979, 3). As Secretary of Defense Caspar Weinberger testified to Congress: "I would rather err on the side of doing too much if that is, indeed, the error, rather than doing to little. It is fire on the side of doing too little" (HASC 1983, 128).

The unanticipated cumulative effect of many "hedges" is a stiffening and enlargement of the requirements for war fighting and deterrence.[28] Hardness numbers were commonly hedged. For example, without giving the evidence for their "hedge" the CBO (1978*b*, 52) used conservative assumptions about the hardness of Soviet indus-

[27] The possibility that one's own actions could be causing defensive reactions by the adversary was rarely explicitly considered, although game theorists, concerned with strategic interaction, do take this into account. Another exception is the discussion in the USA about building new strategic nuclear bombers where the likelihood that the Soviets would have to put resources into air defense systems against bombers (and would therefore not be able to expend vital resources on other, offensive weapons) was used in making arguments about the utility of manned bombers.

[28] While worst case scenarios and hedging because of uncertainty may unconsciously lead to "threat inflation," deliberate threat inflation to justify strategic plans and programs also occurred.

try: "To hedge against Soviet civil defense measures, it is assumed that half of the Soviet industrial base is hardened to 30 psi..." Hedging also applied to the number of potential targets: "For the purpose of estimating force effectiveness...enough weapons are included in a reserve force to maintain effective retaliatory capability, even if there is such a large growth in the number of industrial targets that, by 1990, there would be a 40 percent increase in the number of weapons required to achieve equivalent damage results" (CBO 1978*b*, 51). The public version of the 1994 *Nuclear Posture Review* includes a discussion of a "necessary hedge," although it is not clear how that "necessary" hedge is to be determined (DOD 1994, 12, 14, 16, 18, 19). A graph shows that the "Upload Reconstitution Hedge" accounts for thousands more nuclear warheads in the US stockpile than without the hedge, in order to reconstitute US nuclear forces "should political relations with Russia change for the worse" or there are failures in implementation of the START I and START II arms control treaties (DOD 1994, 14, 19).

5. The Scientific Seduction of Systems Analysis

As noted above, both critics and practitioners of systems analysis raised some of these concerns during the cold war. Practitioners themselves also issued cautionary notes, though such warnings were apparently more common earlier rather than later in the cold war. Sir Solly Zuckerman, an important British nuclear strategist, wrote in 1953 that strategy was "based upon assumptions about human behaviour which seem totally unreal. It neither constitutes scientific analysis nor scientific theorizing, but is a non-science of untestable speculations" (quoted in Freedman 2003, 171). During the 1950s, scholars at the RAND Corporation and elsewhere produced studies emphasizing what they called the "pitfalls" of systems analysis (e.g. Kahn and Mann 1957). Quade (1968*b*, 363), summarizing these pitfalls argued, "No matter how we strive to maintain standards of scientific inquiry or how closely we attempt to follow scientific methods, we cannot turn military systems analysis into an exact science." Enthoven and Smith (1971, 71) wrote that "Some have criticized systems analysis on the grounds that it tends to overemphasize factors that can be reduced to numbers and underemphasize factors that cannot." They grant that this is a "potential danger," that it is "possible for an analyst to become so intrigued with the measurable aspects of the problem that he gives inadequate attention to nonquantitative factors." Yet, they argue that this is "less likely to occur under systems analysis approaches than under alternative approaches" because in using systems analysis "an individual must lay out all his assumptions, objectives, and calculations." Similarly, Charles Hitch (1965, 57) argues that the "systems analyst, like any other scientist, must be prepared to submit his work to critical scrutiny, and not just by other systems analysts. This is one of the great merits of the scientific method—it is an open, explicit, verifiable, and self-correcting process."

But by the 1980s, there was a sense that assumptions and the models themselves need not be examined. Systems analysis was taken to be policy neutral, a sort of "scientific-technical grounding" that was alluded to in congressional hearings on the MX missile by Scowcroft Commission member John Deutsch as "technical examination" by those who were "more technically inclined" which yielded "net technical judgment" (HASC 1983, 101). Thus, technical analysis and modeling was so taken for granted that it was not necessary to produce the figures. One simply had to believe the more technically inclined. Commission Chairman Brent Scowcroft in explaining his belief that 100 MX was the right number, argued: "There is nothing magic about 100. We felt, first of all, that we wanted a number less than that which in conjunction with the other accurate Minuteman force would constitute a first strike against the Soviet Union, their hard targets, their leadership, nuclear storage and so on" (HASC 1983, 86).

Thus, even the cautions described by the first generations of systems analysts appear to have been mostly forgotten by the 1980s as scholars and practitioners sought ways to sharpen the nuclear debate. In their critical overview of nearly two decades of public assessments within the United States of the US–Soviet strategic balance, Salman, Sullivan, and Van Evera argue that "Discourse succeeds when it rests on sound methods of inquiry; the [nuclear] balance debate has failed as a discourse because its methods have been unsound" (1989, 177). Salman, Sullivan, and Van Evera show how flawed analysis can be used to manipulate the political debate and lead to misleading conclusions. They suggested four common games that analysts play: using static indicators or bean counts; flawed dynamic analysis based on bad numbers or faulty assumptions; using outlandish scenarios; and oracle or *ex cathedra* pronouncements by experts making assertions without evidence.

Like others before them who recognized and detailed some of the pitfalls of certain forms of policy modeling, Salman, Sullivan, and Van Evera urge that the solution is better analysis. They argue that "military strength should be assessed by measuring the capacity of forces to execute strategy. . . . using data describing the characteristics of the forces on both sides, the analyst measures the strength of the force by asking whether it can perform its assigned missions, and if so, under what conditions and with what degree of confidence." They suggest that: "To be meaningful, measures of the Soviet–American nuclear balance should describe what both sides' nuclear forces can do. This requires dynamic analysis that assesses their ability to perform wartime missions" (1989, 176). They then use "dynamic analysis" to simulate nuclear exchanges. Their analysis is quite thorough, and to facilitate transparency they provide an appendix discussing the techniques and assumptions of their analysis as well as a computer program so that readers can conduct their own dynamic analysis. They also warn that their analysis should be understood "as an approximation of reality, not a replica. . . . Nuclear war is a mysterious, unprecedented event" (1989, 213). But, they then suggest that their simulations "probably approximate reality as closely as public data will allow, and our interviews suggest that classified simulations produce similar results" (1989, 213). Salman, Sullivan, and Van Evera conclude their article by arguing that rigorous dynamic systems analysis should "define serious" nuclear discourse and determine what gets published:

> Policy concerns will always distort balance assessment to some degree, but scholars of security affairs can mitigate the problem by setting and enforcing higher professional standards. Specifically, they could require that research purporting to measure American nuclear strength, or dealing with issues that require its measurement, provide dynamic analysis that tests the propositions its advances. The provision of such information should define serious work on strategic nuclear issues; manuscripts that omit it should not be published or cited as authority. The academic community can impose such standards if it chooses, and the quality of net assessment will improve if it does. (1989, 244 5)

Thus, even as they document the sloppy use of policy modeling, Salman, Sullivan, and Van Evera simply propose better modeling. They have not, apparently, understood how there was both on the one hand, no way for the modeling to be more accurate, and on the other hand, how the modeling itself began to make the nuclear world.

6. Conclusion: How Abstraction Makes a World

Systems analysis was intended to help policy makers understand the complex and essentially unknown nuclear world and assist them in making the policy process more rational. It was intended to produce usable knowledge, to quantify and model the nuclear world. As Enthoven and Smith (1971, 64) say, "In any analysis, the assumptions drive the conclusions:" the virtue of systems analysis was the ability to use it to explore "all assumptions" and, "In this important sense, systems analysis becomes a method of interrogation and debate suited to complex issues. . . . a set of ground rules for a constructive and divergent debate." But while Enthoven and Smith recognize that assumptions drive the conclusions, they and other users of systems analysis were less than attentive to the ways that systems analysis is not simply analysis. The political-military discourse—in the sense of what we do and don't talk about, and how we talk about it—was structured in subtle and not so subtle ways by systems analysis.

As Enthoven and Smith suggest, "The issue here is not numbers versus adjectives, but clarity of understanding and expression. Numbers are an important part of our language. Where a quantitative matter is being discussed, the greatest clarity of thought is achieved by using numbers, even if only expressed as a range" (1971, 69). Yet as one prominent systems analyst wrote, "Quantification is desirable, but it can be overdone; if we insist on a completely quantitative treatment, we may have to