Stockpile Confidence Under a Nuclear Test Ban

Steve Fetter

A ban on all nuclear testing is one of the oldest and most elusive proposals to control nuclear armaments. Progress toward this goal has been slow and incremental: explosions in the atmosphere, space, and underwater have been banned since 1963 by the Limited Test Ban Treaty (LTBT), underground weapon tests were limited to yields less than 150 kilotons (kt) by the Threshold Test Ban Treaty (TTBT) of 1974, and nuclear explosions for purposes other than weapon testing were similarly constrained by the Peaceful Nuclear Explosion Treaty (PNET) of 1976.[1] Throughout this time, proponents of continued testing have argued that test restrictions inhibit improvements in nuclear warheads that would increase the safety, security, survivability, and effectiveness of weapon systems, and that Soviet observance of a test ban cannot be verified. When the Carter Administration opened negotiations on a comprehensive test ban (CTB) in the fall of 1977, the nuclear weapon establishment in the United States and the United Kingdom raised a new objection: that the reliability of existing nuclear weapons cannot be maintained without nuclear testing. This argument helped unify the bureaucratic opposition, forcing President Carter to place a five-year and then a three-year limit on the length of the proposed treaty. Most recently, stockpile confidence objections have been a central premise in the Reagan Administration's refusal to consider additional restrictions on testing.

The stockpile confidence objection has two parts, both of which will be examined below: (a) that to forego nuclear testing would significantly un-

I would like to thank Carol Alonso, Paul Brown, Ash Carter, Richard Garwin, Charles Glaser, Warren Heckrotte, John Holdren, and J. Carson Mark for their helpful comments.

Steve Fetter is a research fellow at the Center for Science and International Affairs in the Kennedy School of Government, Harvard University. He received an S.B. in physics from the Massachusetts Institute of Technology and a Ph.D. from the University of California, Berkeley. He was a postdoctoral fellow at Lawrence Livermore National Laboratory before coming to CSIA.

1. For a review of the history of the test ban debate, see National Academy of Sciences, *Nuclear Arms Control: Background and Issues* (Washington, DC: National Academy of Sciences Press, 1985), pp. 187–223, Glenn T. Seaborg, *Kennedy, Khrushchev, and the Test Ban* (Berkeley: University of California Press, 1981), and G. Allen Greb, "Comprehensive Test Ban Negotiations, 1958–1986: An Overview" in Jozef Goldblat and David Cox, eds., *Nuclear Weapon Tests: Cessation or Limitation* (Oxford: Oxford University Press, forthcoming).

International Security, Winter 1987/88 (Vol. 12, No. 3)

dermine confidence in the stockpile, and (b) that this would threaten national security by weakening nuclear deterrence. To avoid many of the difficulties of definition that have muddied the debate on this issue, stockpile confidence is here taken to mean a reasonable assurance that the stockpile of nuclear weapons already deployed at the time a CTBT is negotiated would continue to perform, with existing delivery systems, with the degree of reliability necessary for deterrence. Upgrading the stockpile without testing is a separate issue which is addressed elsewhere.[2]

The Confidence Issue vs. the Reliability Issue

Stockpile confidence is not the same as stockpile reliability. Reliability is an objective measure of the average fraction of weapons that will perform properly. Reliability can be measured to any given degree of accuracy by performing a sufficient number of tests. Confidence, on the other hand, is a subjective measure of reliability. Confidence is the belief of those responsible for the stockpile that the weapons are reliable. The amount of testing affects the likelihood that these perceptions are correct.

The difference between confidence and reliability could become extreme in the absence of continued nuclear testing. One could have perfect confidence in weapons that would be unreliable if used, or one could lose confidence in weapons that were perfectly reliable. Therefore, proving that nuclear weapons could be kept reliable during a CTB is not the same as proving that confidence could be maintained. Confidence would be based mostly on scientific experience and non-nuclear techniques, but it would also be subject to political and psychological distortions.

This is not a trivial point, since deterrence is more a matter of perception (confidence) than reality (reliability). If American leaders are convinced of the reliability of their weapons, and Soviet officials, observing this confidence, are also convinced of the potency of the U.S. arsenal, then the requirements of deterrence are satisfied independent of the actual reliability of the weapons. The difference between confidence and reliability would only be revealed on the fateful day that deterrence failed.

The likelihood of problems with complex technologies is chronically underestimated by experts. Nuclear reactors claimed to have unsurpassed reliability

2. Steve Fetter, "The Effect of a Nuclear Test Ban on Weapon Modernization," *Scientific Aspects of Global Security*, Vol. 1, No. 1 (to be published, 1987).

suffer meltdowns. The Space Shuttle, which knowledgeable engineers and program managers claimed was extremely reliable, exploded after two dozen flights. One particularly relevant example of misplaced confidence was the high confidence of U.S. weapon designers in modifications made to stockpiled weapons during the 1958–61 Moratorium on nuclear testing—confidence was so high, in fact, that the modified designs were not tested immediately after the Moratorium ended. One nuclear weapon that was modified during the Moratorium, the W52 warhead for the Sergeant surface-to-surface missile, entered the stockpile in 1962 without further testing. When it was tested in 1963 it gave only a small fraction of its expected yield.[3]

Of course, weapon designers, looking back on this experience, might begin to err in the other direction, losing confidence in reliable weapons during a CTB. But designers could be expected to perform every non-nuclear test possible to determine the reliability of a weapon, and if all the results were positive it would be difficult to conclude that the weapon was unreliable. It is at least as likely that, as with the Space Shuttle, weapons would be assumed reliable unless strong evidence to the contrary was presented. This suggests that those responsible for the stockpile may be more likely to maintain confidence in unreliable weapons than to lose confidence in reliable ones.[4] Although lives may be placed at risk by misplaced confidence in many of the products of complex technologies, such as nuclear reactors, it is not clear whether having confidence in unreliable nuclear weapons would be good or bad, because there are great differences of opinion about what nuclear weapons are supposed to do. Some claim that nuclear weapons should never actually be used and that they would be of little benefit in defending the country that did use them; others maintain just the opposite.

The first section of this paper focuses on the question of whether the nuclear arsenal can be kept sufficiently reliable without nuclear testing. Beliefs about this question will naturally have a strong influence on one's confidence in the stockpile during a CTB. Past instances of unreliability are reviewed, and techniques to detect, assess, and resolve reliability problems without nuclear testing are described. The degree of reliability necessary for nuclear weapons is then discussed. Finally, I return to the subject of the implications of decreased confidence.

---

3. Paul S. Brown, "Nuclear Weapon R&D and the Role of Nuclear Testing," *Energy and Technology Review*, September 1986, p. 13.
4. York has believed that the opposite is more likely. Herbert F. York, letter to Norris Bradbury, Richard Garwin, and J. Carson Mark, September 21, 1978.

Maintaining Reliability without Nuclear Testing

The problem of maintaining nuclear weapons without nuclear testing is a sensitive issue because many of the arguments bear on details of weapon designs and techniques that are classified. The arguments cannot be resolved simply by referring to those who in the past have been responsible for assuring the reliability of nuclear weapons, since they are often divided on this issue. For example, former Los Alamos National Laboratory (LANL) director Norris Bradbury and Lawrence Livermore National Laboratory (LLNL) director Herbert York and former weapon designers J. Carson Mark, Richard Garwin, and Hans Bethe maintain that nuclear testing is not necessary to insure the reliability of the stockpile, while current LLNL director Roger Batzel and former LANL directors Harold Agnew and Donald Kerr contend that testing is essential.[5] Nevertheless, insight into this question can be gained by the careful application of common sense to unclassified historical examples.

PAST EXAMPLES OF UNRELIABILITY

Many problems with nuclear weapons have been detected in the past, not only in the development phase, but also at various times after deployment. Some, but not all, of these problems would have resulted in a greatly reduced yield or no yield at all. Fourteen of the forty-one types of nuclear weapons introduced into the U.S. stockpile after 1958 developed reliability problems; of these problems, 75% were discovered and/or corrected by nuclear testing.[6] The problems these warheads developed, along with two others described

5. See letter to President Jimmy Carter from Norris Bradbury, Richard Garwin, and Carson Mark, August 15, 1978, in House Armed Services Committee Report No. 95–89, "Effects of a Comprehensive Test Ban Treaty on United States National Security Interests," August 14, 15, 1978, Appendix 3, p. 181. Also see letter to Representative Dante Fascell from Hans Bethe, Norris Bradbury, Richard Garwin, Spurgeon M. Kenney, Jr., Wolfgang Panofsky, George Rathjens, Herbert Scoville, Jr., and Paul Warnke, May 14, 1985, and letter to Representative Henry J. Hyde from Roger E. Batzel and Donald M. Kerr, June 7, 1985, both in Hugh E. DeWitt and Gerald E. Marsh, "Weapon Design Policy Impedes Test Ban," *Bulletin of the Atomic Scientists*, Vol. 41, No. 10 (November 1985), pp. 11–13. For Harold Agnew's views, see interview in *Los Alamos Science*, 152 (Summer/Fall 1981), p. 152. The position of Herbert York, as expressed in his letter to Norris Bradbury, Richard Garwin, and J. Carson Mark, September 21, 1978, is somewhat more complicated. York believes that reliability can be maintained, but doubts that the current leadership of the nuclear establishment could maintain confidence in this reliability. He therefore suggests beginning with a nuclear test quota treaty allowing no more than two tests per year yielding no more 10 kt to settle the nerves of the leadership.

6. Paul S. Brown, "Nuclear Weapons R&D and the Role of Nuclear Testing," talk given at the American Physical Society meeting, San Francisco, CA, January 28, 1987.

by J.W. Rosengren,[7] are listed in Table 1. One might be tempted to conclude that about one-quarter of the stockpile would have been unreliable if testing had not been available, but the reliability problems do not lend themselves to such a simple interpretation.[8] The twenty-three problems described in Table 1 can be grouped into six categories: tritium decay, one-point safety, corrosion of fissile material, deterioration of high explosive, low temperature performance, and other problems.

TRITIUM DECAY. During the Moratorium, weapon designers began to suspect that the performance of weapons with aged tritium might have been overestimated. Tritium is radioactive, and decays with a half-life of twelve years. This fact was known before the Moratorium, of course, but weapons were stockpiled without testing for this effect because it was thought to be well understood. When the Moratorium ended, nuclear tests confirmed that a problem existed. Seven nuclear weapons were tested in 1962 to assess and resolve this problem: the B28, B43, and B57 bombs, the W44 warhead for the anti-submarine rocket, the W45 warhead for the Little John, Terrier, and Bullpup tactical missiles and the medium atomic demolition munition, the W50 warhead for the Pershing I missile, and the W59 warhead for the Minuteman I intercontinental ballistic missile (ICBM). More recently, the W84 warhead for the ground-launched cruise missile (GLCM) experienced performance problems related to aged tritium, but presumably this was a more complicated situation (details about this case have not been declassified).

ONE-POINT SAFETY. Nuclear weapons are designed to be "one-point safe," which means that only a trivial nuclear yield will result if the high explosive is detonated accidentally.[9] In the past, four nuclear weapons experienced

7. J.W. Rosengren, "Some Little-Publicized Difficulties With a Nuclear Freeze," RDA–TR–122116–001 (Arlington, VA: R&D Associates, October 1983).

8. Details about problems with the W45, W47, W52, W56, W58, and W68 have been brought out in the on-going debate between Jack Rosengren and Ray Kidder. See J.W. Rosengren, "Some Little-Publicized Difficulties"; Ray E. Kidder, "Evaluation of the 1983 Rosengren Report from the Standpoint of a Comprehensive Test Ban," UCID–20804 (Livermore, CA: Lawrence Livermore National Laboratory, June 1986); Jack W. Rosengren, "Stockpile Reliability and Nuclear Test Bans: A Reply to a Critic's Comments," RDA–TR–138522–001 (Arlington, VA: R&D Associates, November 1986); and Ray E. Kidder, "Stockpile Reliability and Nuclear Test Bans: Response to J.W. Rosengren's Defense of His 1983 Report," UCID–20990 (Livermore, CA: Lawrence Livermore National Laboratory, February 1983). Also see Brown, "Nuclear Weapon R&D," pp. 13–14.

9. The formal definition of one-point safety requires that "in the event of a detonation initiated at any one point in the high-explosive system, the probability of achieving a nuclear yield greater than 4 pounds of TNT equivalent shall not exceed one in one million." Thomas B. Cochran, William M. Arkin, and Milton M. Hoenig, *Nuclear Weapons Databook*, Vol. I, *U.S. Nuclear Forces and Capabilities*, (Cambridge, MA: Ballinger, 1984), p. 67.

Table 1. Nuclear weapon types known to have had stockpile confidence problems[a]

Weapon System		Production Period	Problem	Year Problem Resolved
B28	bomb	1958–66	Improve one-point safety	1962
			Performance with aged tritium	1962
B43	bomb	1961–65	Improve one-point safety	1962
			Performance with aged tritium	1962
W44	ASROC	1961–68	Performance with aged tritium	1962
W45	Little John, Terrier, MADM, and Bullpup	1962–66	Corrosion of fissile material	1963
			Performance with aged tritium	1964
			Deterioration of high explosive	1965
W47	Polaris A1	1960–64	Neutron vulnerability	1962
			Corrosion of fissile material	1963
			Improve one-point safety	1963
W50	Pershing I	1963–65	Performance with aged tritium	1962
W52	Sergeant	1962–66	Improved high explosive safety	1963
W56	Minuteman II	1963–69	Improve one-point safety	1963
B57	ASW bomb	1963–67	Performance with aged tritium	1962
W58	Polaris A3	1964–67	Corrosion	1970s
W59	Minuteman I	1962–63	Performance with aged tritium	1962
B61	bomb	1979–present[b]	Low temperature performance	1981
W68	Poseidon C3	1970–75	Deterioration of high explosive	1980
W79	8-inch AFAP	1981–83	Performance with new gas fill when original reservoir could not be manufactured	1982
W80	ALCM	1981–present	Low temperature performance	1981
W84	GLCM	1983–present	Low temperature performance	1981
			Performance with aged tritium	1984

NOTES:

a. Except for the W56 and W58, these problems are taken from Frederick Reines, Lew Allen, Edwin L. Goldwasser, Andrew J. Goodpaster, Arthur K. Kerman, M. Brian Maple, Kenneth McKay, William G. McMillan, Herbert F. York, and Rochus E. Vogt, "Nuclear Weapons Tests: The Role of the University of California—Department of Energy Laboratories," Report to the President and the Regents of the University of California by the Scientific and Academic Advisory Committee, July, 1987, p. 35. The W56 and W58 are from J. W. Rosengren, "Some Little-Publicized Difficulties with a Nuclear Freeze," RDA-TR-122116-001 (Arlington, VA: R&D Associates, October 1983), pp. 18–20. Production dates are from Thomas B. Cochran, William M. Arkin, Robert S. Norris, and Milton M. Hoenig, *Nuclear Weapons Databook*, Vol. II, *U.S. Nuclear Warhead Production* (Cambridge, MA: Ballinger, 1987), pp. 10–11.

b. Six different modifications of the B61 have been produced since 1966, but the modification that experienced the low-temperature problem has not been specified. Based on the nature of the problem and its timing, the most likely candidates are the B61-3 or the B61-4, which were manufactured beginning in 1979.

problems related to one-point safety: the B28 and B43 bombs, the W47 warhead for the Polaris A1 submarine-launched ballistic missile (SLBM), and the W56 warhead for the Minuteman II ICBM. Of these, only the problems with the W47 and the W56 have been described in the open literature. In the case of the W47, a test just before the Moratorium indicated that the warhead was not inherently one-point safe. Since further testing was not permitted, a mechanical safing device was added to the weapon. With time, however, this safing mechanism jammed. In the first version of the warhead (W47Y1), this problem was resolved without requiring nuclear testing. When the second version (W47Y2) developed similar problems, the decision was made to modify both versions, with the aid of nuclear testing, to make them inherently one-point safe, thus eliminating the need for mechanical safing. The W56 warhead was also developed without enough tests to ensure inherent one-point safety. It too was fitted with a mechanical safing mechanism that jammed after a few years. In this case, however, the problem was eliminated with a small design change that did not require nuclear testing.

CORROSION OF FISSILE MATERIAL. In two cases, the W45 tactical warhead and the W47 Polaris warhead, the normal stockpile surveillance program discovered that the fissile material was corroding. A nuclear test of an extremely corroded W47 indicated that warhead performance was not very sensitive to the effects of corrosion. Extremely corroded units were removed from the stockpile based on this test, and a slight design change prevented further similar deterioration in both warhead types. The W58 warhead for the Polaris A3 SLBM is also known to have developed a corrosion problem. This problem was evaluated by non-nuclear techniques, which indicated that the observed deterioration had little effect on warhead performance. No nuclear tests were done.

DETERIORATION OF HIGH EXPLOSIVE. The stockpile surveillance program discovered dimensional changes in the high explosive of the W45 tactical warhead. This problem was resolved by modifying the high explosive and its packaging, and testing the modified warhead. After several years in the stockpile, the high explosive in the W68 warhead for the Poseidon C3 SLBM also began to deteriorate. Fortunately, the W68 had been tested during development with another high explosive that did not have this problem. The switch was made and a modified warhead was tested successfully.

LOW-TEMPERATURE PERFORMANCE. Paul Brown mentions that one weapon gave a small fraction of its expected yield in a test at the low-temperature extreme of its operating environment.[10] Nuclear tests have been done on three stockpiled weapons—the B61 bomb, the W80 warhead for the air-launched cruise missile (ALCM), and the W84 GLCM warhead—to assess and correct their performance at low temperatures.

OTHER PROBLEMS. Three other problems have been mentioned in the unclassified literature. First, the W47 Polaris warhead was found to be unacceptably vulnerable to neutrons from nearby nuclear explosions. Nuclear tests were done to evaluate solutions to this problem. Second, the W52 warhead for the Sergeant surface-to-surface tactical missile was designed with a very sensitive high explosive. Two fatal accidents occurred during the first production of the weapon, killing four people. A decision was made to change the high explosive, but the Moratorium prevented testing the modified weapon. As mentioned above, scientists were confident that the weapon would perform satisfactorily, but in a nuclear test after the Moratorium it gave only a small fraction of its rated yield. The weapon was then redesigned and tested. Finally, the design of the W79 warhead for the eight-inch artillery-fired atomic projectile (AFAP) had to be modified when a certain component could not be manufactured.

All of the reliability problems described above resulted either from design errors that were not revealed by the normal testing program during the development phase of a weapon or from the unexpected aging and deterioration of weapon components after deployment. This distinction is important, because most of the examples of past failures presented as proof that testing is essential for stockpile reliability in fact reflect an incomplete testing program. Known failures of this type need never be repeated if they are eliminated during development.

The difficulties with one-point safety, for example, are a mixture of design and aging problems. Although it is true that the mechanical safing device in the W47 jammed due to corrosion (an aging problem), this device would not have been required in the first place if enough tests had been done to ensure inherent one-point safety. But, in its haste to deploy new weapons before the Moratorium took effect, the U.S. stockpiled warheads such as the W47 after a truncated development and testing program. The W52 was modified, but not tested because of the Moratorium, to correct flaws discovered during

---

10. Brown, "Nuclear Weapon R&D," p. 14.

the first production of the warhead. Future problems of this type can be prevented by resisting the temptation, should a CTB become a reality, to stockpile the latest designs or modify existing weapons in ways that cannot be certified without nuclear tests.

Other problems occurred simply because weapon designers did not appreciate certain aspects of weapon design or the behavior of the weapon in its stockpile-to-target sequence (STS). The problems with tritium decay, for example, were not due to aging per se, but to a lack of understanding about the effects of this aging on weapon performance. Because of this lack of understanding, nuclear tests were not done during weapon development to assess the consequences of tritium decay. Similarly, because of a lack of theoretical understanding and because non-nuclear tests at low temperature did not produce obviously suspicious results, nuclear tests were not done at the low temperature limit. Little information is available on the neutron vulnerability of the W47, but it is likely that the warhead was not subjected to a nuclear effects test before deployment, for otherwise this problem would have been discovered.

There may be additional shortcomings in design and testing practices that are yet to be identified, but, as with most other technologies, the number of such problems decreases as the technology becomes more mature. Although no testing program can be complete in the sense that all possible design flaws are anticipated, it should be noted that all of the flaws mentioned above were discovered soon after the first production of a warhead. Of the eighteen examples of design flaws mentioned above, six were discovered and corrected in the same year that quantity production of the warhead began, six were resolved one year after first production, three two years later, and one three years later. The two problems that took the longest to identify—the one-point safety and tritium decay problems of the B28—were corrected within four years of first production. By relying only on warheads that have been tested, produced, and deployed for some years, it appears that problems due to design errors can be substantially avoided.

Problems caused by aging and deterioration are more difficult to deal with because they cannot be eliminated by a thorough testing program during the development phase. Unexpected aging and deterioration are more likely to be the major stockpile confidence problem under a prudently-implemented CTB. By their very nature, nuclear weapons must contain reactive materials such as plutonium and high explosive. Warheads are routinely tested after undergoing simulated accelerated effects of aging, but this does not catch all the problems that can crop up during the lifetime of a weapon. Three warheads—the W45, W47, and W58—experienced corrosion problems. None of these problems was detected by nuclear tests, and only one nuclear test was done to evaluate the problems (a successful test of an extremely corroded W47 warhead). In two cases—the W45 and W68—the high explosive deteriorated. Nuclear tests did not detect these problems, although tests were done to certify the modified warheads. In the three cases where mechanical safing devices deteriorated—the W47Y1, W47Y2, and W56—two were resolved without nuclear testing. Therefore, nuclear tests were done to resolve four of the eight cases of unexpected aging and deterioration described in the open literature. All four of these tests were successful. Thus, nuclear tests were not necessary to maintain the actual reliability of the stockpile. They were done to maintain the confidence of those responsible for the stockpile.

Demonstrating that nuclear testing was instrumental in resolving past problems does not prove that testing would be required in the future, however. First, the fact that nuclear tests were *done* does not mean that tests were *necessary*, since the problems were resolved in a context in which nuclear testing was permitted. If testing had not been permitted, confidence might nevertheless have been sufficient in some warheads (e.g., the W68 had already been tested with the substitute high explosive), or alternate approaches might have solved the problem without nuclear testing (e.g., safing mechanisms in the W47Y2 could have been repaired often or a jam-free mechanism developed). Second, most of the past problems in the stockpile appeared around the time of the Moratorium, when the technologies being deployed were new. The key concepts for reliable, high yield-to-weight thermonuclear warheads had just been invented, as had intercontinental and submarine-launched ballistic missiles. It is not surprising that these young technologies exhibited growing pains. But now that the technology of nuclear weapons is more mature and well understood, problems should be less likely in the future. It is sometimes said that new designs are less reliable because they are more complex, but reliability is influenced more by technological maturity and production expertise than by complexity per se.

DETECTING PROBLEMS IN THE STOCKPILE WITHOUT NUCLEAR TESTING

Although some claim that nuclear testing is required to detect reliability problems, this is questionable for two reasons. First, the only reliability problems that were detected by nuclear testing in the past were caused by design flaws, not by aging and deterioration. Except for the W52, nuclear tests were not used to detect even these design problems, but instead have been used to assess recognized possible problems, as in the case of tritium decay. Second, the rate of nuclear testing for these purposes has been very low. In recent years only 8% of all tests were done for stockpile confidence purposes, and few of these were of old weapons.[11] Even if two old warheads were tested every year, this would allow only one stockpile confidence test every fifteen years for each weapon type, since there are twenty-five to thirty-five types of nuclear weapons in the stockpile at any given time. Furthermore, a single successful stockpile confidence test gives little information about the reliability of the weapon, since even a weapon only 50% reliable would give a successful test half of the time. The normal stockpile surveillance program, which consists of the careful dissassembly, inspection, and testing of components from many weapons, is far more effective than nuclear testing for detecting deterioration.

EVALUATING PROBLEMS WITHOUT NUCLEAR TESTING

Doing a nuclear test is often the simplest way to tell if a deteriorated warhead will perform properly. Even so, testing for these purposes is relatively rare: only eight of nearly 300 tests since 1970 were done to evaluate defects in stockpiled weapons.[12] There are many tools and techniques other than nuclear testing that can be used to assess the reliability of old weapons. First, non-nuclear testing of weapons can be done by substituting inert materials for the special nuclear materials (SNM) and then detonating the high explosive. This type of test would be useful in evaluating the aging of a high explosive, for example. Small amounts of fissile material could be used in the experiment, as was done during the Moratorium, resulting in a very small nuclear yield (less than a pound of high explosive equivalent).[13] Second, the operation of nuclear weapons is routinely simulated and studied with elaborate computer programs running on the fastest computers in the world. These programs have been adjusted to give correct results for past nuclear tests, and may therefore be more valuable in assessing the performance of

11. Robert S. Norris, Thomas B. Cochran, and William M. Arkin, "Known U.S. Nuclear Tests: July 1945 to 16 October 1986," Nuclear Weapons Databook Working Paper 86–2 (Washington, DC: Natural Resources Defense Council, October 1986), p. 10.
12. Ibid., p. 11.
13. Robert N. Thorn and Donald R. Westervelt, "Hydronuclear Experiments," LA–10902–MS (Los Alamos: Los Alamos National Laboratory, February, 1987).

old weapons than in predicting the performance of new warheads (which is currently their most common use). Computer simulations might determine, for example, whether a slight change in the composition or shape of certain components would affect the yield of a weapon. Although these techniques have not spotted every problem in the past, they were instrumental in solving problems with many warheads. Moreover, the power of these techniques has increased consistently with the passage of time. Even though they cannot by themselves guarantee proper operation of a weapon, past experience gives a good indication of the range of problems for which they can be trusted to give reliable advice.

CORRECTING PROBLEMS WITHOUT NUCLEAR TESTING

Laboratory techniques are best at indicating whether deterioration problems are serious and at suggesting minor changes that have been verified by testing experience. To solve serious reliability problems without nuclear testing one must either remanufacture the warhead or find an acceptable substitute. These approaches are discussed below.

REMANUFACTURING. If deterioration occurs and laboratory procedures indicate that the weapon may no longer be reliable, the weapon could be remanufactured to the original specifications. Only one U.S. warhead has been remanufactured: the production line was closed for three years in the early 1970s, and the remanufactured weapon was not proof-tested.[14] Remanufacturing has been so rare because weapons are usually replaced by newer designs before the end of their lifetime. The main objection to remanufacturing is that it rebuilds old problems into new warheads: designing new weapons might solve aging and deterioration problems, improve warhead safety and military effectiveness, and save money. But if the advantages of a CTB outweigh the marginal benefits of these modernizations, and if the warhead to be rebuilt had a reasonably long shelf life, then remanufacturing can be an effective solution. Three difficulties with replicating nuclear weapons have been raised: materials, fabrication techniques, equipment, and workmanship quality change in subtle ways over time, with unpredictable results for weapon performance; the production process requires trained weapon designers, who will become less and less skilled under a CTB; and

14. Paul S. Brown, Lawrence Livermore National Laboratory, personal communication, October 13, 1986.

remanufacturing old warheads is less cost-effective than building new, state-of-the-art warheads.[15] These are not insoluble problems.

1. The difficulty of remanufacturing nuclear warheads has been compared with that of rebuilding a 1950s television set: even if one had the schematic diagram, where would one find the proper parts? In the U.S., warheads typically are manufactured during a five- to ten-year period, after which the production line is closed. Since nuclear weapons are designed to have a shelf life of about twenty-five years, fifteen to twenty years could elapse between the original manufacture of the warhead and its remanufacture. Although this might be the case with warheads that are already out of production when a CTBT takes effect, the production lines for warheads currently under production could be kept open indefinitely to minimize the possibility of changes in materials and techniques. At present, the B61, W76, W80, B83, and W87 strategic nuclear weapons are being produced, and in the next few years the W88 and the warhead for the new short-range attack missile (SRAM II) will be produced (see Table 2). Furthermore, warheads that are out of production but are planned to be part of the long-term stockpile, such as the W78, could be brought back into production before a CTB takes effect.

Rebuilding the 1950s television might not be difficult if, during the original manufacture, one had taken special precautions to minimize the problems of replication. The composition and structure of materials can be specified with extreme accuracy, as can the techniques used to produce a material. Chemists sometimes complain that they don't know all the variables to specify in order to replicate some substances, or the deviations that can be permitted without affecting weapon performance, but studies can and should be undertaken to resolve these issues. If questions remain about the possibility of duplicating a material, it could be produced without interruption or stockpiled. Fabrication techniques can also be specified with great precision using the computer-aided design and engineering systems that are already in use at these facilities.[16] Warhead components from the original production line could be stored for direct comparison with remanufactured components. And while it is often argued that tightening occupational health and safety standards could prevent the industrial use of certain key materials, this need

15. Letter from Batzel and Kerr in DeWitt and Marsh, "Weapon Design Policy," pp. 12–13. Also see letter from Frank J. Gaffney to Representative Edward J. Markey, January 21, 1986, in Hugh E. DeWitt and Gerald E. Marsh, "An Update on the Test Ban," *Bulletin of the Atomic Scientists*, Vol. 42. No. 4 (April 1986), pp. 10–12.
16. William H. Hubbell, Jr., "The Weaponization Program," *Energy and Technology Review*, September 1986, p. 33.

Table 2. Bomb and warhead types in the U.S. strategic arsenal of the late 1980s and early

	Bomb/warhead type	Production[a] Period
W56	Mk-11C Minuteman II warhead	1963–69
B61	Gravity bomb	1966–present
W62	Mk-12 Minuteman III warhead	1970–76
W68	Mk-3 Poseidon C-3 warhead	1970–75
W69	SRAM warhead[b]	1972–76
W76	Mk-4 Trident C-4 and D-5 warhead	1978–present
W78	Mk-12A Minuteman III warhead	1979–82
W80	ALCM and ACM warhead	1981–present
B83	Gravity bomb	1983–present
W87	Mk-21 MX and Midgetman warhead	1986–present
W88	Mk-5 Trident D-5 warhead	Not yet produced

NOTES:
a. Production dates are from Robert S. Norris, Thomas B. Cochran, William M. Arkin, and Milton M. Hoenig, *Nuclear Weapons Databook*, Vol. II, *U.S. Nuclear Warhead Production* (Cambridge, Mass.: Ballinger, 1987).
b. To be replaced by the SRAM II in the mid-1990s.

not prevent the use of the same materials in weapons; the cost of additional protection for personnel should not be an issue if national security were at stake.

Admiral Sylvester R. Foley, then Assistant Secretary of Energy for Defense Programs, agreed with this analysis in 1986: "Assuming, therefore, that vendor-supplied materials and components are still available at the time desired for remanufacture (and this will not necessarily be the case), the remanufacture of existing, well-tested warheads is possible."[17] Assuring the supply of the required materials should not be an impossible task for a country as large and advanced as the United States.

Even so, many claim that nuclear tests of remanufactured weapons would be required since nothing complex can be relied on without testing. But some systems, such as nuclear artillery or communication satellites, are considered sufficiently reliable without critical nuclear tests. The U.S. has not fired a live modern nuclear artillery shell because above-ground tests are banned and

17. Admiral Sylvester R. Foley, responses to questions for Department of Energy budget hearing before the Subcommittee on Procurement and Military Nuclear Systems, Armed Services Committee, U.S. House of Representatives, February 19, 1986.

underground tests would be too expensive. The new and expensive MILSTAR and NAVSTAR satellite systems are intended to provide critical command and control communications during a nuclear war, even though they cannot be tested under these conditions. If there is no crisis of confidence about systems that have not been thoroughly tested, why should one lose confidence in thoroughly-tested weapon designs that have been remanufactured?

2. The stock of weapon scientists with nuclear testing experience would diminish gradually under a CTB, and the cessation of testing and weapon development could accelerate the loss of trained scientists and make the recruitment of high-quality personnel more difficult. There are several interrelated questions.

First, would it be difficult for the weapon laboratories to keep experienced scientists and hire high-quality personnel under a CTB? If weapon scientists are motivated solely by the desire to design and test weapons that will go into the stockpile, then a CTB would lead to a complete loss of such personnel. But individual motivations are much more complex, and include the desire to discover and publish, freedom to think creatively, recognition by peers, access to state-of-the-art equipment, and the desire to contribute to the national security. The fact that weapon laboratories would no longer be designing new devices for the stockpile would undoubtedly lead some people to leave, but many motivations for weapon-related work would not change significantly under a CTB. Much work would remain that is challenging and creative, laboratory equipment could still be first-rate, and the contribution to the national defense just as important. Scientists wanting a new challenge could move to non-weapon programs at weapon laboratories, where they would still be available for consultation about stockpile problems.

Second, even if reasonably good scientists were available, could they keep the stockpile reliable without the skills and practical experience that nuclear testing can give? Although tests are essential to confirm predictions about designs that extend the state of the art, they are not necessary to maintain established designs. Production of a warhead requires experienced weapon designers, but only in the initial phases when it is uncertain whether the production processes can match the specifications of the designer. After this, there is little designer involvement. Furthermore, every activity other than nuclear testing that contributes to design expertise would be available under a CTB. Besides exploring the theoretical aspects of weapon design, weapon scientists could investigate many aspects of weapon physics by using the non-nuclear testing and computer simulation techniques mentioned above, or by improving the ability of computer simulations to predict the behavior of stockpiled weapons for comparison with existing test data. Experiments in a wide variety of areas could be done using the small fusion explosions created in the laboratory with inertial confinement fusion (ICF). ICF would train technical people in the unique disciplines of direct relevance to nuclear weapons, and would retain and attract talent to the weapon laboratories.[18] Even though direct experience in nuclear testing is unlikely to be important for maintaining the stockpile, there would still be scientists with testing experience twenty years after a CTBT was negotiated.

Third, would those responsible for maintaining the stockpile have confidence in their own work without recourse to nuclear testing? Would decision-makers have confidence in the judgment of these people? These questions underscore the difference between confidence and reliability. As pointed out above, well-trained, competent people who make correct decisions could lose confidence, and incompetent people could nevertheless be completely self-confident. Many of today's weapon designers say that they would be much less self-confident without access to nuclear testing, even if they were only responsible for maintaining old designs. This may turn out to be true, but it is hard to find examples of a similar loss of self-confidence in other technical fields. Those responsible for the stockpile during the Moratorium were self-confident, and those responsible for the satellite systems mentioned above do not express doubts (if they have any) about the functioning of those systems as intended in an environment they have never been tested in. There may be a perverse effect at work here: without experimental data to prove them wrong, scientists become more confident in their theoretical judgement. This effect might be exaggerated if those who are most comfortable with experimental proof leave the laboratories.

It is not clear that the actions of decision-makers would depend much on the self-confidence of scientists or the adequacy of their training. Decision-makers are usually in no position to judge the adequacy of a scientist's training or the quality of his judgment. If doubts about the stockpile were expressed, decision-makers might believe that the scientists were being overly cautious, or that the risks of a problem were not outweighed by the benefits of ignoring it. This was the case with the Challenger disaster, where the worries of engineers were overridden by program managers concerned about long delays. A technical community, even one as exclusive as the

---

18. Roger Batzel, letter to Admiral Sylvester R. Foley, Jr., July 22, 1986.

nuclear weapon community, is rarely unanimous in its advice. It is likely that decision-makers will be influenced primarily by political considerations, and that they will choose the technical analysis they want to hear. For example, decision makers who were against a CTBT in the first place might use a perceived technical problem to justify a decision to abrogate the treaty. On the other hand, those who do not want to upset the international political equation may choose to ignore the problem.

Fourth, it is often said that a CTB would lull the United States into a false sense of security, and that budgetary support for the weapon laboratories—whose primary mission, the development and testing of new weapons, would have been effectively banned—would evaporate. In the past, however, excessive fear of a lulling effect has resulted in its opposite, stimulating actions that reduced the value of the treaties signed.[19] It is unlikely that a CTBT would be ratified by the U.S. Senate without considerable support from the defense community, and especially from the Joint Chiefs of Staff. This support, in turn, is unlikely without certain assurances, or safeguards, that would partially compensate for the loss of nuclear testing. In the case of the LTBT, these safeguards included well-funded weapon laboratories, a vigorous underground testing program, and maintenance of the capability to resume atmospheric testing should the Soviets violate the treaty.[20] In the case of a CTBT, the safeguards might include excellent funding for laboratories, a vigorous non-nuclear testing program, and maintenance of the capability to resume underground testing at short notice. It is true that the U.S. has allowed its ability to promptly resume atmospheric testing to lapse ten years after the treaty was signed. Some claim that this is evidence that safeguards cannot be maintained, but the U.S. has only let this safeguard lapse because of the realization that nearly all important tests can be done underground (which was not thought to be true in 1963), and that prompt resumption of atmospheric testing is not critical to protect national security if the Soviets abrogate the LTBT.

Finally, it has been argued that the Soviet Union would develop a relative advantage over the United States in the quality of weapon scientists since it

19. Ivo H. Daalder, "The Limited Test Ban Treaty," in Albert Carnesale and Richard N. Haass, eds., *Superpower Arms Control: Setting the Record Straight* (Cambridge, MA: Ballinger Publishing Company, 1987). Also Sean M. Lynn-Jones, "Lulling and Stimulating Effects of Arms Control," ibid.

20. U.S. Congress, Senate, Committee on Foreign Relations, "Nuclear Test Ban Treaty, Hearings on Executive M," 88th Congress, First Session, 1963, "Letter to Senator Russell from the Chairman of the Joint Chiefs of Staff re: safeguards recommended by the JCS," p. 982.

can force trained scientists to continue in weapon-related work. But creative thinking cannot be forced, and Soviet scientists would also live under the constraint of a test ban. One may just as well argue that U.S. scientists would have an advantage, because U.S. tests probably have been performed with better instrumentation and diagnostics, and because the computers available for weapon simulations are vastly superior to those in the Soviet Union. Hence, the loss of trained personnel does not favor the Soviet Union if both parties abide by the treaty. But if the Soviet Union did test clandestinely at low yields and at a small but steady rate, this could give them an advantage in expertise if the treaty collapsed at a later date.

3. There will be additional costs incurred if remanufacturing is necessary to restore confidence in aging weapons that would have remained in the stockpile had testing been available. Although modern nuclear weapons are designed to have a shelf life of about twenty-five years, a shorter lifetime might be more appropriate under a CTB. A stable stockpile might consist of 15,000 to 20,000 weapons of fifteen to twenty different types. If the weapon lifetime was ten years, for example, 1,500 to 2,000 weapons would be manufactured each year, or an average of 100 weapons of each type. This rate of manufacture is well below the current capacity of 3,500 to 4,000 warheads per year.[21] Assuming that an average nuclear weapon costs roughly one million dollars, this remanufacturing program would cost less than two billion dollars per year, or less than one-third the amount now spent on nuclear weapon research, development, testing, and production activities.[22] Funds may also be required to assure the availability of commercially-produced materials, either by subsidizing the vendor or by government production, but since most of the cost of a warhead is probably accounted for by special nuclear material (SNM), fabrication, and assembly costs, that is unlikely to be important.[23]

---

21. Thomas B. Cochran, William M. Arkin, Robert S. Norris, and Milton M. Hoenig, *Nuclear Weapons Databook*, Vol. II, *U.S. Nuclear Warhead Production* (Cambridge, MA: Ballinger, 1987), p. 12.

22. Nuclear warheads cost 10 to 15% of total system costs. Brown, "Nuclear Weapon R&D," p. 7. Total system costs per deployed warhead (in current dollars) range from $6 million for Minuteman III to $13 million for Trident. Cochran, *U.S. Nuclear Forces*, pp. 119–140. The projected 1987 expenditures for weapon research, development, testing, and production are $4.4 billion, and another $2.6 billion will be spent on weapon materials production and waste management. Cochran, *U.S. Nuclear Warhead Production*, p. 21.

23. There are also "virtual costs" of remanufacture, because one could have used the same money to buy more effective, custom-designed warheads for new delivery vehicles. From the viewpoint of CTB proponents, however, this is one of the main advantages of test ban: to the degree that nuclear testing is important in developing "more effective" (i.e., counterforce) weapons, a ban on testing would have a stabilizing effect on the arms race in general.

INTERCHANGEABLE WARHEADS. If laboratory techniques and remanufacturing are not sufficient to restore confidence in an aged warhead, there is often an alternate warhead that could be used on the delivery vehicle. Table 2 lists the warheads that will constitute the U.S. strategic arsenal during the late 1980s and early 1990s. It is not usually recognized that many warheads have a natural backup already in the stockpile. Consider, for example, the ICBM force. If the new and sophisticated W87 warhead used in the Mk-21 reentry vehicle (RV) for the MX (and probably Midgetman) missile should develop serious stockpile confidence problems, then the Minuteman III W78 warhead/Mk-12A RV combination could be substituted. The MX has been successfully flight-tested with Mk-12As.[24] Even though the W78 warhead lacks the insensitive high explosive and more sophisticated fusing mechanisms of the W87, the military effectiveness of the MX would not be significantly affected by this change. The decrease in accuracy of the MX missile resulting from use of the Mk-12A warhead would reduce the probability of destroying hard targets by less than 10%.[25] This hardly represents a fatal flaw in deterrence, unless one believes that deterrence is weak now because the MX is not yet

24. William M. Arkin, Andrew S. Burrows, Richard W. Fieldhouse, Thomas B. Cochran, Robert S. Norris, and Jeffrey I. Sands, "Nuclear Weapons," in Frank Blackaby, ed., *World Armaments and Disarmament: SIPRI Yearbook 1985* (London: Taylor and Francis, 1985), p. 45.

25. The accuracy of ballistic missile reentry vehicles is primarily determined by the inertial guidance system of the missile, the fusing system of the RV, and the aerodynamic properties of the RV. The Mk-21 and Mk-12A RVs have slightly different aerodynamic properties, arising from differences in the ballistic coefficient (a measure of atmospheric drag) and the symmetry of the ablation of the RV. A discussion of guidance errors is found in Matthew Bunn, "Technology of Ballistic Missile Reentry Vehicles," in Kosta Tsipis and Penny Janeway, eds., *Review of U.S. Military Research and Development: 1984* (Washington, D.C.: Pergamon-Brassey's, 1984), pp. 67–116. Inertial guidance errors limit accuracy to a CEP of no less than about 30 m. D.G. Hoag, "Ballistic-missile Guidance," in B.T. Feld, T. Greenwood, G.W. Rathjens, and S. Weinberg, eds., *Impact of New Technologies on the Arms Race* (Cambridge, Mass.: M.I.T. Press, 1971, p. 81. Assuming that this has been achieved in the guidance system of the MX missile, and that the CEP of the MX is about 100 m, the CEP due only to reentry errors would be 95 m (the total error is the square root of the sum of the squares of the individual errors). Assuming a ballistic coefficient of 3,000 lb/ft^2 for the Mk-21 and 2,200 lb/ft^2 for the Mk-12A (Matthew Bunn, "U.S. Ballistic Missile Reentry Vehicles: A Research Note," unpublished, August 1986), then, according to the figure on page 71 of Bunn, the reentry error of the Mk-12A is about 1.3 times greater than that of the Mk-21. The CEP of the MX using the Mk-12A would therefore be roughly 25 m greater than when using the Mk-21. Assuming a 300 kt yield for the Mk-21 and a 335 kt yield for the Mk-12A, the probability of destroying a target hardened to 2,000 psi would be decreased by about 5% by using the Mk-12A instead of the Mk-21 on the MX. Even if the CEP was 40 m greater, the kill probability would only be decreased by 9% (assuming a reliability of 85%); for a two-on-one attack, the decrease would only be 3%. The accuracy of the Mk-12A probably could be improved by using advanced fusing and nose-tip technology developed for the Mk-21.

fully deployed. Similarly, if the W78/Mk-12A failed, one could use the other Minuteman III warhead/RV, the W62/Mk-12, which is functionally identical to the W78/Mk-12A except for having only half the yield. One could even guard against undiscovered failures by initially deploying a mix of warheads on each missile type, so that the complete failure of a particular warhead would not disable an entire missile system.

Similar arguments can be made for the other legs of the triad. Consideration of the SLBM force is somewhat more complicated because of the anticipated reliance (in the long term) on a single missile—the Trident II—and the fact that details about the warhead payload have not been made public. One author speculates that 75% of the Trident II missiles will be armed with eight of the new W88/Mk-5s now under development, with the remainder carrying twelve to fourteen of the older W76/Mk-4s.[26] These warheads could substitute for one another, although the yield of the W76 is only about one-fifth that of the W88. Against soft targets, fourteen W76s would be about 60% as effective as eight W88s in causing blast damage. Against hard targets, fourteen W76s would be about as effective as eight W88s,[27] although avoiding fratricide effects would complicate targeting. One could also consider using ICBM warheads on Trident II. Since the W78 and W87 were considered as warheads for the Trident II, it is likely that they could also substitute for the W88 (and vice-versa).[28]

In the case of the bomber force, the B61 and B83 gravity bombs would be natural substitutes for each other; if confidence lessened in one, greater reliance could be placed on the other. The short-range attack missile (SRAM), which uses the W69 warhead, is soon to be replaced by the SRAM II with a new warhead. If this new warhead develops problems, one could resort to the old warhead and missile. The same cannot be said for the cruise missiles, since both the ALCM and the new advanced cruise missile (ACM) being developed use the W80 warhead. It may be possible to use the W84 GLCM warhead or the SRAM II warhead with the ALCM since they are deployed in similar circumstances and have about the same yield as the W80.

---

26. Robert S. Norris, "Counterforce at Sea: The Trident II Missile," *Arms Control Today*, September 1985, p. 7.
27. Reentry errors for SLBM RVs are less than for ICBM RVs due to the greater angle of reentry. Even if the W76/Mk-4 had a CEP 15% greater than the W88/Mk-5, the hard-target kill capability of the two systems would be equal, assuming fourteen W76s with 100 kt yield, eight W88s with 475 kt yield and a CEP of 100 m, a Trident II missile reliability of 85%, and a target hardness of 2,000 psi.
28. Cochran, et al., *U.S. Nuclear Forces*, p. 145.

Although many of these substitutions should be straightforward, it is conceivable that a nuclear test would be desired in some cases. This possibility should be investigated, and, if necessary, nuclear tests performed, *before* a CTB takes effect. Warheads could be tested for the accelerations, vibrations, and temperature extremes that might be expected with other delivery vehicles. The delivery vehicles may have to be flight-tested with the substitute warheads, but a CTB does not prevent this. Such a program would undoubtedly be expensive, but it is an option for those who believe that, for unforeseen reasons, remanufacture may not always work.

SUPER-RELIABLE WARHEADS. Modern U.S. warheads are designed close to many technological edges: for a given task, the overall volume or mass of the weapon or amount of SNM is minimized. Many test failures during the development of a weapon occur because the yield of the primary (or fission component) is not large enough to cause the secondary (or fusion component) to ignite. Since the secondary accounts for most of the yield of a weapon, the resulting failure is nearly total. In many cases, this is overcome by adding more SNM to the primary, which increases the cost of the warhead.

If priorities were changed so that weapon designers strove for maximum reliability instead of the most "bang for the buck," it is likely that weapons could be made robust than is the case at present. An active program to design super-reliable warheads that are as far from known technological edges as possible, are easy to remanufacture, and which use the most deterioration-resistant materials available could be begun before a CTB takes effect. These warheads could be designed to substitute directly for critical warheads in the stockpile, such as the W80, B61 or B83, and W87 or W88. The super-reliable warhead need not be larger or heavier than the original, but it may cost considerably more (because of the increased use of SNM), the yield may be lower (because of the greater volume of high explosives), or one may have to give up certain safety features (insensitive high explosive or relaxation of the one-point safety criterion). The development and testing of such weapons probably could be completed in about five years. Note that these warheads need not be stockpiled in large quantities; the design could simply be an insurance policy against potential catastrophic failures in the regular stockpile.

If such a development program is deemed too costly, there are still many things that could be done short of developing entirely new warheads. For example, each stockpiled warhead could be tested with two or three different kinds of high explosive, so that if the material of choice deteriorated, the problem could be solved without nuclear testing, as was the case with the W68. Alternate materials could be developed for each critical weapon component, or tests done for the effects of extreme deterioration. Experienced weapon designers could doubtless think of many more things that could be done to prepare for a CTB if time and money were allocated for this purpose.

Warhead Performance in Context

Weapon reliability is often stated in the most simplistic terms: either every weapon of a given type gives its full rated yield, or every weapon is a dud. Although an entire stockpile of duds could conceivably result from serious design or manufacturing errors, as happened with the W52 warhead, it is an unlikely consequence of weapon deterioration. Indeed, most of the deterioration problems in the past would not have resulted in a stockpile of duds for two reasons. First, deterioration, which is by nature a random process, affects only a fraction of the stockpile at a given time. Only a fraction of the mechanical safing devices in the W47 and W56 jammed; only a fraction of the W45, W47, and W58 warheads experienced severe corrosion of the fissile material; and only a fraction of the W58 and W68 warheads experienced deterioration of the high explosive. It may take years after a problem is first spotted for a significant fraction of the stockpile to be affected, and this time can be used to formulate an acceptable solution to the problem. Second, many aging problems result in a reduced yield, not a complete failure. Corrosion resulted in only a slightly reduced yield for the W47, and an aged W45 gave half its rated yield. Many military missions can still be accomplished with a reduced yield.

RELIABILITY

Current levels of testing are insufficient to determine the actual reliability of a particular type of weapon with a high degree of certainty. Even if ten confidence tests were performed and all were successful, there would still be a 30% chance that the weapon would be less than 90% reliable, and a 10% chance that it would be less than 80% reliable.[29] To be 95% sure that the

29. The probability of m successes out of n tests of a weapon with reliability r is $C_{n,m}r^m(1 - r)^{n-m}$, where $C_{n,m} = n![m!(n - m)!]^{-1}$. If all the tests are successful (m = n), and all reliabilities have the same a priori probability, then the probability P that $r \geq R$ is $(1 - R^{n+1})$, and the number of tests required for given values of P and R is $\log(1 - P)/\log(R) - 1$. Weapon designers implicitly assume a very low a priori probability for values of the reliability other than zero or one, so for

weapon is at least 95% reliable would require nearly 60 tests, all successful. This number of tests is, of course, out of the question and has never been performed. One could argue that a small number of tests for individual weapon types is sufficient because all weapons have many things in common, but this is only true if the failure modes of weapons due to deterioration are similar, predictable, and well-understood, which is not the case.

In spite of this statistical reasoning, some weapon designers continue to insist that nuclear weapons are either 0% or 100% reliable. This may be true at the theoretical design level, but it is unlikely to be true of aging effects. The simultaneous failure of every unit in use is unheard of in commercial technologies. Imagine all blenders failing at the same time because a rubber gasket cracks, or every light bulb burning out after the same amount of use. It is difficult to understand why nuclear weapons should be thought to exhibit this pattern of catastrophic failure due to aging. Indeed, it appears paradoxical for weapon designers to claim on the one hand that nuclear weapons are such delicate devices that seemingly harmless variations in the manufacturing process can have remarkable effects on weapon performance, and yet on the other hand maintain that weapons are so uniform in behavior that they all could fail completely at the same time.

Weapon designers insist that nuclear weapons should be extremely reliable—nearly 100% reliable—even though this very high degree of reliability is excessive when it is put in the context of the many other much greater uncertainties that exist when using nuclear forces. Consider the events that must occur to use an ICBM effectively: (a) the National Command Authority (NCA) must issue the command; (b) the command must be received by the launch crew; (c) the launch crew must execute the order; (d) the missile and its guidance system must perform correctly; and (e) the nuclear warhead must explode. The reliability of each of the first four steps is far from perfect. Missile reliability is commonly assumed to be about 75 to 90%.[30] No one

---

them a single test is enough to tell whether a warhead type is perfect or a dud. The reason most often given for this assumption is that warhead performance is highly non-linear (i.e., small changes in a certain parameter can induce large changes in yield). But since modern warheads are designed close to the technological edge, and since the variations induced by aging are random and are often in the parameter space where the yield varies rapidly, a reliability between zero and one is quite possible. The only way to find out for sure is to test.

30. Some argue that weapon reliability is more important than missile reliability because launch failures can be instantly compensated for, while weapon failures cannot. In fact, retargeting reserve missiles to execute the attack plans of launch failures is a time-consuming process. In principle, one could compensate for weapon (or guidance) failures by having the warhead radio back that the mission was a success (see Richard Garwin, "Bombs that Squeak: Nuclear Explosion

knows what fraction of the launch crews actually would obey orders. The communications systems—especially satellite systems—used to transmit the order have never been tested under wartime conditions, in which nuclear weapons may be exploding at the rate of ten per second. Even the survival of the NCA, not to mention the missiles themselves, is not certain.

Seen in this context, a warhead reliability of, for example, 98% greatly exceeds reasonable system requirements. There is no logical reason to insist on such a high reliability. There should be little cause for immediate alarm if one finds that 5% or 10% of a given type of weapon are showing signs of deterioration. This level of unreliability has a small effect on military effectiveness, and several years are likely to elapse before uncertainties about the unreliability of the warhead overwhelm the large uncertainties in the rest of the military system. The reliability of the MX warhead could decrease by 25% and still have the same probability of destroying hard targets as a Minuteman III warhead.[31] One should also remember that no single warhead design bears the whole weight of deterrence. There are currently eleven major strategic nuclear warheads spread among three classes of delivery vehicles: the SLBM force with three warhead types and the ICBM and bomber forces with four types each. The most numerous strategic warhead planned, the W80, will account for less than 20% of the total.[32]

YIELD

The uncertainties in yield measurements are typically about 10%, and a nuclear test generally is considered successful if the estimated yield falls within 10% of the predicted yield.[33] Many tests have failed, but few tests have resulted in no yield at all. This is because the primary almost always detonates even though its yield may not be great enough to ignite the secondary. Similarly, aging and deterioration problems seldom result in zero yield. In the case of the W45 warhead for the Little John missile, the stockpile confidence test gave half the rated yield. Although this was considered an

---

Location During Nuclear War," unpublished, November 1981), or by monitoring the targets with nuclear burst detection systems aboard satellites.

31. This calculation assumes that the MX has a yield of 300 kt and a CEP of 0.05 nmi, and Minuteman III has a yield of 335 kt and a CEP of 0.10 nmi (Cochran, *U.S. Nuclear Forces*, pp. 116–121), and that both are used against a target hardened to 2,000 psi (a typical silo).

32. The number of ALCMs and ACMs planned is 3,150. The total number of strategic weapons is about 16,000. Robert S. Norris and William M. Arkin, "Nuclear Notebook," *Bulletin of the Atomic Scientists*, Vol. 43, No. 5 (June, 1987), p. 56.

33. Paul S. Brown, personal communication.

awful failure by weapon scientists, the military effectiveness of the weapon would not have been seriously affected by this reduction in yield.

Consider the effect that reductions in yield have on the probability of destroying hardened military targets such as missile silos or command bunkers. The probability that the target will be destroyed depends mainly on four variables: the yield and the accuracy of the weapon, the hardness of the target, and the combined reliability of the command, missile, guidance, and warhead systems. A simplified equation for the probability of kill is given by

$$P_k = R\{1 - 2^{-(EY/CEP^3H)^{2/3}}\}$$

where Y is the yield in megatons (Mt), CEP (Circle of Equal Probability) is the accuracy in nautical miles (nmi), H is the hardness in pounds per square inch (psi), R is the system reliability, and E is a constant that determines the overpressure at a given distance from the explosion (for surface bursts and high over-pressures, E is about 16.4 nmi^3psi/Mt).[34] If more than one weapon is used against the target, then the probability of kill is

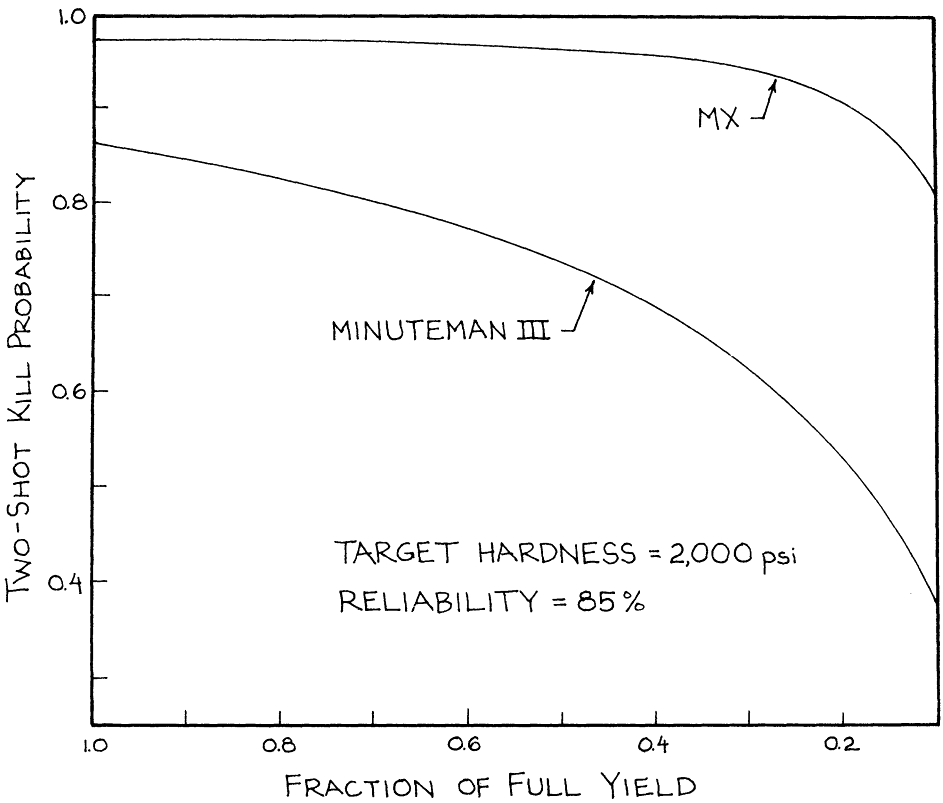
$$P_k(n) = 1 - [1 - P_k]^n$$

where n is the number of weapons. (This assumes that the warheads are delivered by different missiles, that their reliabilities are independent, and that there are no fratricide effects.)

Figure 1 shows the decrease in the kill probability as a function of warhead yield for a two-on-one attack of MX and Minuteman missile warheads against a target hardened to 2,000 psi (a typical silo). The MX is assumed to have a CEP of 0.05 nmi and a yield of 300 kt, and Minuteman to have a CEP of 0.1 nmi and a yield of 335 kt.[35] Both missiles are assumed to be 85% reliable. Note that decreases in yield have very little effect on the kill probability of the MX. Even if the yield of the MX warhead drops by 85%, the kill probability would be the same as that of Minuteman III. The results for the Midgetman and Trident II missiles should be similar to those for the MX since they are expected to have approximately the same yield and accuracy.

34. Kosta Tsipis, *Arsenal: Understanding Weapons in the Nuclear Age* (New York: Simon and Schuster, 1983), pp. 305–308. More complicated formulae take into account the height of the burst and the fact that the hardness of hard structures is better described by the maximum impulse they can withstand rather than the maximum overpressure, but these differences are unimportant for the examples given in the text.
35. Cochran, *Nuclear Weapons Databook*, pp. 75, 118, 121, 126.

Figure 1. The decrease in kill probability as a function of warhead yield.



Unless one is prepared to argue that counterforce capability on the order provided by the MX, and not much less, is essential for deterrence, then even large yield reductions do not matter for accurate warheads.

Modest increases in accuracy and overall system reliability can compensate for decreases in yield. For example, a 21% decrease in CEP or a 4% increase in system reliability would compensate for a 50% reduction in the yield of the MX warhead. If the weapon yield were only one-tenth of its expected value, this could be offset by halving the CEP. Another consideration is "bias," or the possibility that the actual wartime flights of ballistic missile RVs might be deflected from their targets because of unanticipated systematic errors. Former Secretary of Defense James Schlesinger stated that a bias of 0.1 to 0.2 nmi might be possible for the Minuteman III system;[36] if this is true, then reducing the bias by half would compensate for a 20 to 60% reduction in yield of that system.[37]

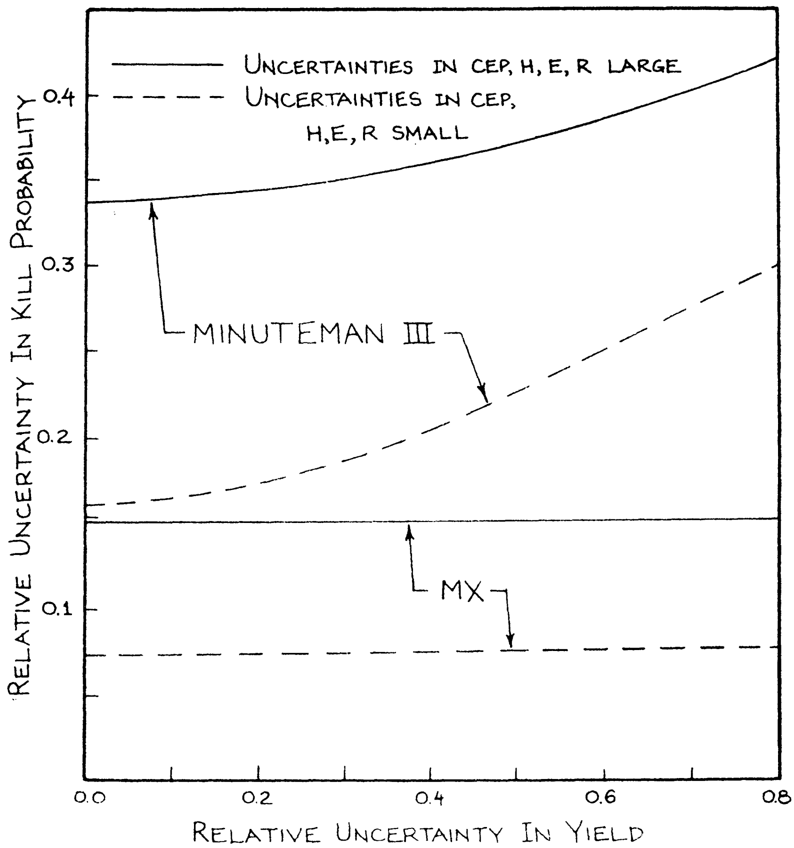
Stockpile confidence problems caused by aging may simply result in a greater uncertainty about yield rather than a certainty of reduction. Figure 2 shows the increase in the relative uncertainty in the kill probability as a function of the relative uncertainty in weapon yield for the MX and Minuteman III systems. Note that uncertainties in yield have no effect on the uncertainty in kill probability for the MX, and that large yield uncertainties can be tolerated for the Minuteman III. For example, increasing yield uncertainties from 10% to 50% increases the uncertainty in kill probability by only 0.3–1% for MX and 10–39% for Minuteman, depending on the size of uncertainties in yield, CEP, hardness, overpressure, and reliability.

There is already great doubt that some missions—such as destroying a large percentage of an opponent's ICBM force—can be carried out at all, and absolute reductions in yield or increases in uncertainty about yield are unlikely to change that situation much. Many aging problems would result in the expectation of a reduction in yield, or increased yield uncertainty, but would not result in a total weapon failure. For attacks against moderately hard targets with accurate warheads, yield reductions of at least 50% can be tolerated. Moreover, the uncertainty in the yield can be many times greater than that allowed now without significantly increasing the uncertainty in the kill probability.

For attacks against urban-industrial targets, yields reduced by 95% would still cause tremendous devastation. Although many find the idea of countervalue attacks repugnant, or implausible since they would invite a similar response, the possibility of such attacks is the foundation of virtually all formulations of deterrence. A force of only 1,000 strategic warheads (7% of the current force) each with a reliability of only 10% and a yield of only 10 kt (1–10% of current strategic warhead yields) could kill at least ten million

36. In testimony before the Arms Control Subcommittee of the Senate Foreign Relations Committee on March 4, 1974, Schlesinger gave examples of 0.1 to 0.2 nmi for the Minuteman missile, which is 100 to 200% of the CEP. Quoted in Matthew Bunn and Kosta Tsipis, "Ballistic Missile Guidance and Technical Uncertainties of Countersilo Attacks," PSTIS Report No. 9 (Cambridge, Mass.: Program in Science and Technology for International Security, Massachusetts Institute of Technology, August 1983), p. 66.
37. The effect of bias is calculated from data in Bunn and Tsipis, "Ballistic Missile Guidance," p. 63.

Figure 2. The increase in the relative uncertainty in kill probability as a function of yield uncertainty.[a]



---

a. The uncertainty in kill probability, σ_P, is given by

$$\sigma_P^2 = \sigma_R^2(\partial P/\partial R)^2 + \sigma_C^2(\partial P/\partial C)^2 + \sigma_H^2(\partial P/\partial H)^2 + \sigma_Y^2(\partial P/\partial Y)^2 + \sigma_E^2(\partial P/\partial E)^2$$

where σ_R, σ_C, σ_H, σ_Y, and σ_E are the uncertainties in reliability, CEP, hardness, yield, and overpressure. Substituting for $(\partial P/\partial R)$, $(\partial P/\partial C)$, $(\partial P/\partial H)$, $(\partial P/\partial Y)$, and $(\partial P/\partial E)$, which are the partial derivatives of the equation for the kill probability with respect to reliability, CEP, hardness, yield, and overpressure, we have

$$\delta_P = a[9\delta_C^2 + \delta_H^2 + \delta_Y^2 + \delta_E^2 + a^2\delta_R^2]^{1/2}$$

where δ_P, δ_C, δ_H, δ_Y, δ_E, and δ_R are the fractional uncertainties in kill probability, CEP, hardness, yield, overpressure, and reliability, and $a=-2f/3(e^{-f}-1)$, where $f=-\ln(2)(EY/H)^{2/3}CEP^{-2}$. Using the value of E given in the text and assuming a target hardness of 2,000 psi gives $a = 0.022$ for the MX ($Y = 0.30$ Mt, CEP $= 0.05$ nmi) and $a = 0.31$ for the Minuteman III ($Y = 0.34$ Mt, CEP $= 0.10$ nmi). The reliability is given by $R = R_c R_m R_w$, where R_c is the reliability of C^3 system, R_m is the reliability of the missile, and R_w is the reliability of the warhead. The uncertainty in the reliability is given by $\delta_R = (\delta_{Rc}^2 + \delta_{Rm}^2 + \delta_{Rw}^2)^{1/2}$.

For the curve labeled "other uncertainties small," $\delta_C = 0.1$, $\delta_H = 0.2$, $\delta_E = 0.2$, $R_c = R_m = 0.95$, $R_w = 0.98$, $\delta_{Rc} = \delta_{Rm} = 0.05$, and $\delta_{Rw} = 0.02$; for the curve labeled "other uncertainties large," $\delta_C = 0.2$, $\delta_H = 0.4$, $\delta_E = 0.5$, $\delta_R = 0.15$, $R_c = R_m = 0.90$, $R_w = 0.98$, $\delta_{Rc} = \delta_{Rm} = 0.10$, and $\delta_{Rw} = 0.02$. The fractional uncertainty in kill probability for two RVs/target is equal to $\delta_P(1 - P)/(1 - P/2)$.

people.[38] "Before, however, we become too bemused with megatons and multimegatons," Norris Bradbury said in testimony to Congress in 1963, "I would urge one to look again at the pictures of Hiroshima and Nagasaki in 1945 after 15 or 20 kilotons—kilotons, not megatons."[39]

Is a Decrease in Confidence Necessarily Bad?

The degree of confidence required in the stockpile depends on how one plans to use nuclear weapons. Advocates of limited nuclear use, escalation dominance, or the extended, warfighting, or war-winning variants of deterrence demand more confidence in the stockpile than do those who espouse minimum deterrence, simply because they envision that nuclear weapons may actually be used at some future date to achieve a rational goal of national policy. In this view, demonstrations of military might, such as nuclear tests, are also essential to maintain the psychology of deterrence—that is, to make the opponent more fearful that nuclear weapons would be used in a conflict, thereby deterring actions that lead to conflict in the first place. Advocates of minimum deterrence, on the other hand, generally feel that the destructiveness of nuclear weapons is so great and that control over their use during war is so difficult that their use can serve no rational goal. They also feel that brandishing weapons is more likely to provoke hostility rather than to deter conflict. If the specter of turning Washington or Moscow into a radiating ruin is sufficient for deterrence, then deliberations about stockpile confidence are mostly irrelevant. This section makes some general observations about stockpile confidence without attempting to take sides on the more fundamental issue of what deterrence is all about.

GRADUAL EROSION OF CONFIDENCE

Many talk about a decline in stockpile confidence in much the same way as one would talk about a decline in reliability: a gradual erosion that takes

---

38. The bomb dropped on Hiroshima had a yield of 12.5 kt (Cochran, *U.S. Nuclear Forces*, p. 32), and an effective lethal area of 8.5 km^2. Samuel Glasstone and Philip J. Dolan, eds., *The Effects of Nuclear Weapons* (Washington, D.C.: Department of Defense and U.S. Department of Energy, 1977), p. 544. For comparison, about 7.2 million people live in the city of New York, which has a land area of 786 km^2. The use of 100 10 kt warheads to destroy the most densely-populated areas in the U.S. would result in at least 10 million deaths. See William H. Daugherty, Barbara G. Levi, and Frank H. von Hippel, "The Consequences of 'Limited' Nuclear Attacks on the United States," *International Security*, Vol. 10, No. 4 (Spring 1986), pp. 3–45.
39. U.S. Congress, "Nuclear Test Ban Treaty," p. 580.

place over years or decades. But there are many reasons for believing that no matter what happened to warhead reliability, confidence would remain high until a serious case of unreliability was discovered, after which confidence would plummet. A catastrophic decline in confidence has happened several times with other technologies. For the sake of argument, however, let us first assume that a gradual decline is possible. Estimates of warhead reliability might decline from essentially 100% to, say, 90%, then 80%, then 70% over a number of years. What effect would this gradual erosion of confidence have on deterrence, crisis stability, and arms race stability?

At one extreme are those who argue that a gradual decline would threaten the very foundation of U.S. nuclear strategy, since the U.S. depends on nuclear weapons for more than just simple deterrence. They argue, for example, that declining confidence would weaken extended deterrence because it would reduce the credibility of nuclear threats in situations short of general nuclear war. As the credibility of nuclear threats diminishes, both countries may be more likely to escalate crises or engage in conventional warfare. This, in turn, may raise the risks of nuclear war, as the losing conventional forces turn to nuclear weapons. Although there is some truth to this argument, it is overstated, for it assumes that the current arsenals—and not much less—are required to deter a Soviet attack. This argument also ignores the fact that battlefield nuclear weapons are important primarily because they raise the probability of strategic nuclear war, not because they could achieve a tactical goal.[40] Tactical nuclear weapons much less reliable than current forces are sufficient to pose the risk of escalation, and strategic weapons much less reliable than current forces are sufficient to pose the risk of devastating retaliation. It is difficult to believe that nuclear deterrence is so fragile that a gradual loss of stockpile confidence would make either country more willing to start a large conventional war or less determined to avoid a nuclear war.

Related to this is the concern that declining stockpile confidence would weaken the security guaranties of allies and compel each to strengthen its own forces, conventional or nuclear. This is less than convincing, for how could other countries know more about the U.S. stockpile than the U.S. would know? The U.S. simply would not allow the stockpile to deteriorate badly enough to worry allies or embolden other countries; corrective action

40. Thomas C. Schelling, *Arms and Influence* (New Haven: Yale University Press, 1966), pp. 107–116.

would be taken long before this point was reached. Furthermore, it is highly unlikely that the stockpiles of the superpowers would degrade so much that the significance of the arsenals of a third party would increase markedly as a result.

Regarding arms race stability, the superpowers might increase the size of their nuclear arsenals in the wake of a CTB to compensate for a perceived decrease in the reliability of their weapons. Although this would be a temptation, vertical proliferation could be constrained by SALT-like treaties, which most CTB proponents hope would accompany a CTBT, that limit the numbers of weapons. Even if these efforts failed and the number of warheads did increase, this does not necessarily mean that the arms race would become unstable or the risk of war higher. Many feel that the increasing military effectiveness of new weapon systems, made possible by continued nuclear testing, creates a greater danger in the arms race than an increase in the number of weapons per se.

At the other extreme are those who try to make declining confidence into a virtue. The main argument here is that declining U.S. stockpile confidence would reduce the incentives for the U.S. to launch a preemptive strike during a crisis, since it would make a successful first strike more difficult to achieve. Similarly, crisis stability would be enhanced if Soviet confidence in their stockpile eroded. Ironically, a similar argument is made by members of the Reagan Administration regarding the SDI: even a partially-effective defense would be desirable because it would hopelessly complicate a first strike. The main difference is that for a CTB the uncertainties would be accepted voluntarily, and no arms race or countermeasures (short of violating or abrogating the treaty) could change the symmetry of the situation, but for SDI the uncertainties are forced on one by the other side, and a race for countermeasures could quickly alter the situation. Moreover, while an opposing defense may create anxiety about the effectiveness of a second-strike force, it is difficult to see how modest decreases in warhead reliability would have this effect.

Even if the argument in favor of decreasing confidence is sound, it would be unwise from the standpoint of U.S. domestic politics to stress this point. As mentioned above, a CTB is only possible with the support (or at least without the organized opposition) of the military, and they are unlikely to be convinced that unreliable equipment is somehow better than reliable equipment. Nevertheless, there is an important connection between decreasing stockpile confidence and the goals of a CTB that should not be missed:

the less confident one is that nuclear weaponry will work as predicted, the less one will depend on it for limited military objectives; the less confident one is in the design of new weapons, the less valuable new weapons become. Most CTB proponents *want* the superpowers to become less dependent on nuclear weapons for keeping the peace. Whether decreased reliance on nuclear weapons will enhance or degrade the national security or make war more or less likely is, of course, debatable.

Some argue that the Soviets would be more confident in their stockpile than the U.S. would be in its stockpile even if the Soviets did not cheat. If this were true, some of the effects of confidence erosion would not be felt by the Soviets, which might leave the U.S. at a relative disadvantage. It has often been pointed out that the Soviets are more cautious and conservative in developing technologies than Americans. Their designs tend to advance in an evolutionary way rather than by exploiting the latest improvements in technology. Supported by evidence that the yield-to-weight ratio of Soviet weapons is less than that of U.S. weapons[41] (which can permit more robust designs), some have argued that their designs would be more reliable under a CTB. Still, the stability of deterrence rests largely on perceptions. If the Soviets do not design near the technological edge it is probably due to simple lack of confidence, not out of a conscious desire to build super-reliable warheads. The Soviet Union, respecting the technological prowess of the United States, may fear that U.S. designs are more reliable than Soviet designs. This is the case in satellite technology, where the Soviets, for all their technological caution and lack of sophistication, still have equipment that is far less reliable.

CATASTROPHIC WEAPON FAILURE

Suppose that, instead of a gradual decline of confidence, the worst happened: an ingredient in the arsenal that was judged to be essential for deterrence deteriorated, and nothing short of nuclear testing could make it right again. The weapon had been diagnosed to have a reliability or yield insufficient to achieve its military mission. Laboratory testing and computer modeling were unable to support an adequate modification that could be trusted without testing. Remanufacturing was for some reason problematic. An acceptable substitute could not be found, or the substitute experienced reliability prob-

---

41. Lynn R. Sykes and Dan M. Davis, "The Yields of Soviet Strategic Weapons," *Scientific American*, Vol. 256, No. 1 (January 1987).

lems as well. A super-reliable replacement warhead had not been designed. What would happen then?

Theoretically, deterrence would only be harmed if the Soviets shared the U.S. perception that the stockpile was unreliable. But how could the Soviets ever be sure that American confidence in the U.S. stockpile had really declined? Laboratory directors might testify before Congress to this effect, but would the Soviets believe it? Or would the Soviets fear that it was merely a ploy to resume testing? Even if they were sure that American confidence had declined, how could they know that the actual reliability of the weapons had decreased, when even Americans would not know this for sure? Furthermore, how could the Soviets extract any military or political benefit from an American lack of confidence?

No matter what the Soviets thought, and no matter what the theoretical implications for the stability of deterrence, the hysteria about a "confidence gap" *would* adversely affect the American-Soviet relationship. If the Joint Chiefs of Staff joined with the weapon laboratories in stating that the deterrent force was crippled, the president would be under tremendous pressure to withdraw from the treaty. Even though a CTBT would almost certainly contain an escape clause that allowed a party to withdraw if its vital interests were threatened, the political ramifications of abrogation would be enormous.

The adverse consequences of a confidence gap are great, but the risk depends on the probability that the scenario will occur, which the foregoing analysis suggests is small. To demonstrate how unlikely this scenario is, let us assign hypothetical probabilities to each failure that must occur before a catastrophic failure results. Based on the past 30 years' experience, the probability that a warhead will experience a serious deterioration problem (i.e., one that would result in an ineffective weapon) is about 10%.[42] The probability that the problem could not be resolved without nuclear testing might be 25 to 50%.[43] For the sake of argument, assume a 10 to 20% probability

---

42. As mentioned earlier, most of the past stockpile problems resulted from design flaws that were detected within a year or two of first warhead production. By relying on warheads that were first produced five years before the start of a CTB, these problems will be substantially eliminated. Only five of the warheads listed in Table 1 experienced aging and deterioration problems: the W45, W47, W56, W58 and the W68. This represents 12% of the forty-one warheads deployed since the Moratorium. The problem rate in newer warheads can be expected to be lower, since the mechanical safing devices and materials-compatibility problems so troublesome in the past have been eliminated.

43. Seventy-five percent of past stockpile problems used nuclear tests for their resolution, but many of these problems were design flaws (see footnote 42). Four of the eight warhead aging

that remanufacturing the warhead is for some reason difficult or impossible. The likelihood that a suitable substitute could not be found, or that the substitute failed as well, might be 10 to 20%. Also assume that the U.S. has not been prudent enough to design a super-reliable warhead, or even to test substitutes for the deteriorating component. Using these conservative assumptions, the probability that one of the eight critical strategic warheads (B61, W76, W78, W80, B83, W87, W88, and SRAM II) would experience a problem during a 20–30 year period that could not be resolved without nuclear testing would be no more than one percent. If two warhead types must fail catastrophically for a crisis of confidence to occur, then the odds would be very much smaller. To be a net advantage to the U.S., the advantages of a CTB must outweigh this small risk.[44]

NUCLEAR TEST QUOTA: THE SOLUTION? As a final measure to remove even the small risk of catastrophic failure, one could allow infrequent proof-tests of stockpiled weapons. One test every year or two might be enough for stockpile confidence, but not for significant modernization. Even if every allowed test were used for the development of new nuclear weapons instead of the proof-testing of old weapons, this low rate of testing would greatly increase the time necessary to develop a new and significantly different type of weapon, which requires about ten tests. Exotic nuclear weapons such as the X-ray laser would require many more than ten tests before even a development decision could be made. Thus, insurance against the small risk of catastrophic failure might be gained while still constraining modernization.

There is one glaring technical problem with a test quota: how does one know how many weapons have been exploded during a single event? It has usually been assumed that multiple explosions would be allowed under a test quota treaty, but this could allow modernization. If, for example, the total yield of the allowed event was 150 kt, then ten 15 kt devices could be detonated simultaneously, thus allowing the development of new nuclear weapons. Weapon development would be more awkward than is now the case, as one could not simply iterate the process of testing and modification until the final design is derived. One would instead have to try many different ideas at once, pick the one that worked best, and try several modifications of that design next year. In this way, it should be possible, though awkward

problems described here (50%) were resolved without nuclear tests. If nuclear testing had not been available, a greater fraction of these problems would have been resolved without testing.

44. This simple analysis ignores common-mode failures (i.e., one problem that would affect several warhead designs at once), but very little can be said about this without a detailed examination of specific warhead design features, which, of course, cannot be done here.

and slow, to develop even advanced weapons such as an X-ray laser, since several directed-energy experiments can be done with a single nuclear test.

One solution to this problem might be to lower the maximum yield of such tests to the lowest yield consistent with verifying stockpile reliability. Since the primary is the most unreliable part of a nuclear weapon, this might be about 15 kt.[45] Muffling the seismic signals to allow greater yields could be prevented by allowing on-site inspection of the test, such as is specified in the Peaceful Nuclear Explosion Treaty (PNET). The number of devices being tested might be determined without gathering sensitive data. A treaty of this kind would be at least as detailed as the PNET, which took 18 months of intense negotiations to conclude. Many of the political advantages of a test ban may be lost if the ban is not truly comprehensive, however, unless a quota is just a short-lived stepping-stone to a CTBT.

Conclusions

Although there are valid reasons for worrying about the reliability of the nuclear stockpile under a CTB, there are many methods other than nuclear testing by which the confidence of those responsible for the stockpile can be maintained. Regular inspections would provide assurance that problems do not exist. If a weapon is found to have deteriorated, extensive laboratory techniques and computer simulations can assess the consequences of the problem and possibly suggest minor modifications. If this is not satisfactory, and if the weapon has enjoyed a reasonably long lifetime, the weapon can simply be remanufactured to its original specifications. If this proves difficult, adequate substitutes exist for most warheads. Super-reliable warheads could be designed and tested before a CTBT is negotiated as an insurance policy against catastrophic failures. Lastly, reduced weapon reliability or yield, or increased uncertainty in reliability or yield, resulting from stockpile problems may not significantly affect the military effectiveness of the weapon system or of the strategic deterrent as a whole.

If estimates of weapon reliability decrease only slowly, this does not appear to harm deterrence, crisis stability, or arms race stability. If the Soviets abided by the treaty, there is no reason that significant asymmetries would develop. But if the U.S. were to suffer a catastrophic failure of a critical weapon system,

45. Frank von Hippel, Harold A. Feiveson, and Christopher E. Paine, "A Low-Threshold Test Ban," *International Security*, Vol. 12, No. 2 (Fall 1987), p. 135.

the president would be under great pressure to withdraw from the treaty—an action which would have enormous domestic and foreign policy implications. Fortunately, the probability of a catastrophic failure is very small, but it is a danger that must weigh in the balance between the risks and benefits of a CTBT. Allowing one 15 kt test every one or two years would nearly remove this risk, but at the expense of making the test ban less than comprehensive and sacrificing much of its political rationale.

To date, the nuclear weapon establishment in the United States has been far more imaginative about the problems a CTB might pose than about possible solutions to these problems. This seems to be changing somewhat, since in 1982 the military characteristics (MCs, or specifications for nuclear warheads) were expanded to call for maximizing warhead lifetime, the ability to replicate the warhead at a future date, and the ability to incorporate the warhead into other delivery systems.[46] Even though these developments are reassuring, one might legitimately ask why these considerations were not included during the more than two decades this country held a CTB as a goal of national policy, and why they are not accorded a higher priority now. If stockpile confidence is a real concern of policy makers, a vigorous program to solve these problems should be started well before a CTB is negotiated. Since the Reagan Administration maintains that a CTB is a long-term national goal, now would be an especially appropriate time to begin planning for one.

---

46. Brown, "Nuclear Weapons R&D," p. 8.