

An Exchange on Stockpile Confidence

John D. Immele; Paul S. Brown; Steve Fetter

International Security, Vol. 13, No. 1. (Summer, 1988), pp. 196-215.

Stable URL:

http://links.jstor.org/sici?sici=0162-2889%28198822%2913%3A1%3C196%3AAEOSC%3E2.0.CO%3B2-1

International Security is currently published by The MIT Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/mitpress.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

Correspondence | John D. Immele and

An Exchange on Stockpile Confidence

Paul S. Brown

Steve Fetter

$\mathbf{T}_{ ext{o}}$ the Editors:

We have read with great interest the article by Steve Fetter recently published in International Security.1 We are concerned with the fundamental thesis of Fetter and others who believe that technical and political confidence in the reliability of U.S. weapons can be high during a comprehensive test ban (CTB). We believe that nuclear weapons are very much like other types of high-technology military equipment tanks, rockets, airplanes, radar systems, etc. However, while these other military systems are extensively tested and retested during development and after deployment, nuclear weapons are certified to function properly over a wide range of stressful conditions on the basis of only a handful of nuclear tests. History has shown that continued nuclear testing has been necessary to maintain the reliability of the nuclear stockpile.

We have found that Fetter's summary of U.S. stockpile experience is misleading in many respects. While he agrees at first that "the likelihood of problems with complex technologies is chronically underestimated,"2 he later discounts the likelihood and importance of problems in the case of nuclear weapons. In several instances, he describes design strategies that could provide a degree of insurance against a CTB. In fact, these strategies have long been part of our design approach. The University

The work of Immele and Brown was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract W-7405-Eng-148.

John D. Immele is responsible for the physics divisions at Lawrence Livermore National Laboratory that design fission, fusion, and nuclear directed-energy weapons. He previously managed activities such as the Laboratory's advanced conventional ordnance effort and the Peacekeeper (MX) warhead development. Paul S. Brown is the Livermore Laboratory's Assistant Associate Director for Arms Control and was a delegate to the U.S.-USSR Nuclear Test Experts Meetings in Geneva, 1986. He has directed development of computer codes to simulate nuclear explosions and to model explosive interactions in geologic media.

Steve Fetter is a research fellow at the Center for Science and International Affairs, Kennedy School of Government, Harvard University. He is the author of Toward a Comprehensive Test Ban, recently published by Ballinger. 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.

2. Ibid., p. 133.

^{1.} Steve Fetter, "Stockpile Confidence Under a Nuclear Test Ban," International Security, Vol. 12, No. 3 (Winter 1987/88), pp. 132–167.

of California Scientific and Academic Advisory Committee (SAAC), comprising senior scientists and defense experts, recently reviewed the nuclear testing practices at the Livermore and Los Alamos Laboratories. They concluded that "the Laboratories have been acting under a plan that emphasizes the necessity to be prepared for a possible CTB."3

There are a number of reasons for continued nuclear testing. The most important is the need to be able to develop new weapons in response to advances in nonnuclear technology and changes in political strategy. Major emphasis is also placed on the process of modernizing nuclear weapons to improve their safety, security, and survivability. However, Fetter's article and our commentary concentrate on stockpile confidence and the need for nuclear testing. Some of the points mentioned in this commentary are elaborated in a recently released, unclassified report to Congress.4 The reader may also wish to consult the Congressional Research Service study⁵ for a thorough survey of all of the test ban issues, pro and con.

The question we address here is whether weapon confidence can keep pace with the requirements of U.S. nuclear deterrence. This will be an important issue for as many years as it takes to achieve a massive reduction in the numbers of nuclear weapons and a decrease in our dependence on them for national security.

It seems to us that the burden of proof in the matter of confidence and nuclear testing belongs to those who favor a test ban. The consequences of being wrong are monumental. Other technologies of this scale have only been successfully and confidently preserved with extensive testing. Ironically, some automatically consider a CTB a good idea until proven otherwise. Our own situation when responding to this idealistic approach reminds us of a similar difficulty recently faced by Space Shuttle scientists when they had to decide whether or not to fly Challenger. Morton-Thiokol engineers expressed reservations about the reliability of the solid rocket booster, but their concern about possible mission failure was ultimately rejected because the shuttle had worked before and because there were economic and schedule pressures to launch:

MR. BOISJOLY (MORTON-THIOKOL ENGINEER): I felt personally that management was under a lot of pressure to launch and that they made a very tough decision, but I didn't agree with it. One of my colleagues that was in the meeting summed it up best. This was a meeting where the determination was to launch, and it was up to us to prove beyond a shadow of a doubt that it was not safe to do so. This is in total reverse to what the position usually is.

^{3.} The University of California Scientific and Academic Advisory Committee (SAAC); Lew Allen, E.L. Goldwasser, A.J. Goodpaster, A.K. Kerman, M.B. Maple, K. McKay, W.G. McMillan, F. Reines, H.F. York, and R.E. Vogt, Nuclear Weapons Tests—The Role of the University of California— Department of Energy Laboratories, A Report to the President and the Regents of the University of California (July 1987), p. 2.

^{4.} George H. Miller, Paul S. Brown and Carol T. Alonso, Report to Congress on Stockpile Reliability, Weapon Remanufacture and the Role of Nuclear Testing, UCRL-53822 (October 1987).

5. Congressional Research Service, "Nuclear Weapons and Security: The Effect of Alternative

Test Ban Treaties," forthcoming.

DR. FEYNMAN (MEMBER OF THE COMMISSION): I take it you were trying to find proof that the seal would fail?

MR. BOISJOLY: Yes.

DR. FEYNMAN: And of course, you didn't, you couldn't, because five of them didn't, and if you had proved that they would have all failed, you would have found yourself incorrect because five of them didn't fail.

In Finding 4 from Chapter V, "The Contributing Cause of the Accident":

The Commission concluded that the Thiokol Management reversed its position and recommended the launch of 51-L, at the urging of Marshall [Flight Center] and contrary to the views of its engineers in order to accommodate a major customer.⁶

In fact, the similarities between reliability of solid rocket boosters⁷ and of nuclear explosives are striking:

- —Both technologies have been perceived by the public and the media to be fairly mature and even reasonably well understood;
- -Both, in fact, have advanced steadily;
- —In both cases, numerical modeling of the high-temperature, high-pressure phenomena are only approximate and usually semi-empirical;
- —In both cases, some less than "full-up" testing (e.g., hydrodynamic tests of fission triggers and rocket test stands) provides valuable insights and reduces, but does not replace, the need for larger full-scale tests;
- —Both types of hardware can fail to perform their intended mission due to subtle changes in either the physics or operation (e.g., temperature, density, and pressure of reacting fuel) or engineering defects (e.g., the O-ring);
- —Both must work the first time and cannot be tested nondestructively; thus they are exhaustively checked and rechecked by inspecting every fitting, chemically analyzing every part, etc.

Nuclear explosives are more complicated than rocket engines. They contain more parts; their engineering is at least as complex; the physics involved is even further from the realm of laboratory experience.

We believe that few technologists or political leaders would propose relying on a military force of solid rocket motors without full-up testing and without a cadre of experts. This seems particularly obvious after the Challenger, Titan, and Delta mission failures. This same logic argues that U.S. nuclear forces must be supported by nuclear testing.

We submit that, in attempting to prove that stockpile confidence can be sustained in a CTB, Fetter has erred in his analysis of a number of key questions:

^{6.} Report of the Presidential Commission on the Space Shuttle Challenger Accident (Washington, D.C.: U.S. Government Printing Office [U.S.G.P.O.], June 6, 1986), pp. 93, 104.

^{7.} In referring to solid rocket boosters, we include such components as the mechanical structure, propellant and igniter, and directional nozzles, but not inertial sensors and guidance computer.

ARE NONNUCLEAR TESTS AND COMPUTER SIMULATION ADEQUATE TO ENSURE WEAPON RELIABILITY AND CONFIDENCE OVER SEVERAL DECADES? In addressing corrosion and aging problems with four weapons (the W47, W45, W58, and W68), Fetter contends that "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."8 Fetter distinguishes between stockpile confidence and reliability, saying that "confidence is the belief of those responsible for the stockpile that the weapons are reliable."9 On this second point, we agree. Weapon experts must have confidence if our leaders are to have confidence. Any loss of confidence would become apparent in our society, and the "perception of deterrence," as Fetter describes it, would suffer. On the other hand, in the closed Soviet society, a similar loss of confidence could be hidden from view.

While it is true that final proof tests of these fixes were successful, Fetter's statement regarding the reliability of the four weapons has the advantage of hindsight. Corrosion was observed in components of the W47 Polaris warhead and the W45 tactical munition. A nuclear test of the W47 was required to evaluate the worst allowable case of corrosion. W47s with corrosion worse than this case were removed from the stockpile because they were judged to be unreliable. The W45 had two additional aging problems; one was only discovered in a nuclear test and the other involved deterioration of the high explosive (HE). The problems were fixed and the redesigned weapons were certified in nuclear tests. Routine stockpile surveillance discovered serious HE degradation in the W68 Poseidon warhead. Fortuitously, we knew from a development test that the replacement HE was nearly equivalent. However, a confidence test was done to restore political and military confidence in a critical leg of the triad. Furthermore, the test surprised us when the yield was lower than expected and this resulted in a necessary modification to the maintenance procedure for the weapon. The assessment of W58 corrosion depended critically on a series of highly diagnosed tests conducted on similar designs; the diagnostics used in these tests were developed long after the W58 was deployed.

Since these weapons generally performed well in proof tests after they were repaired, does that mean the tests were not necessary? We find such reasoning similar to the rationalization for launching Challenger because it had worked five times before, despite engineering concerns about the O-ring. Similarly, we suspect the modified shuttle SRB will also be "reliable" without a proof flight test. However, there will be no "confidence" until after a test flight of the new configuration.

Fetter advocates that we place our confidence in the "many tools and techniques other than nuclear testing," such as nonnuclear tests and computer simulations. The W80 warhead with insensitive HE (IHE) recently failed a low-temperature nuclear proof test, although a nonnuclear hydrodynamic test at the same low temperature generated no serious concern. Indeed, one reason the problem with cold IHE is so often cited is that it exemplifies how nonnuclear tests and computer simulations are

^{8.} Fetter, "Stockpile Confidence," p. 141.

^{9.} Ibid., p. 133.

not sufficient, even in the 1980s, to predict all device failure modes. Consider non-nuclear testing of fission triggers. In a hydrodynamic test (no fissile isotopes) or a hydronuclear test (less than a few pounds TNT equivalent because of reduced fissile isotopes), there is no nuclear explosion. While these tests are very valuable to the development of a weapon, they address only approximately the first step in the primary explosion process. The crucial steps involved with nuclear energy production do not take place in these tests. Such crucial processes as nuclear-driven disassembly, primary boosting and yield production, energy coupling to the secondary, and mainstage (secondary) thermonuclear explosion can only be addressed in nuclear explosion tests.

U.S. weapons depend heavily on a thermonuclear process in the primary called boosting. Boosting very efficiently increases yield-to-weight ratio, but it can only be certified reliably with full-scale testing. Our experience has shown that boosting is a subtle process and that small, seemingly insignificant changes can cause it to fail; thus nuclear testing is essential. Fetter's comments on how "computer simulations might determine . . . whether a slight change . . . would affect the yield of the weapon" of strike us—and probably anyone with practical experience—as naive. These tools certainly failed to predict the malfunction of the IHE primary at cold temperatures. In fact, they fail to predict many of the warhead parameters which we measure; that is one of the reasons we conduct nuclear tests.

It will be a long time, if ever, before computer simulations can replace nuclear testing because there are important phenomena that are incompletely or poorly modeled. The computer models are only as good as our understanding of the physics that is programmed into them. In particular, the way energy flows in dense hot material and the turbulent mixing of burning fuel are critical to weapons design but are still not well understood from first principles. New, more powerful supercomputers can increase the value of our testing and give us more capability and options for design. Increased understanding of and confidence in warhead performance require continual iteration between theory and nuclear experiment.

COULD PAST STOCKPILE PROBLEMS HAVE BEEN AVOIDED IF WE HAD "MORE COMPLETE TESTING" DURING DEVELOPMENT? CAN PROBLEMS DURING A CTB BE AVOIDED BY RELYING ONLY ON WARHEADS THAT HAVE BEEN TESTED AND DEPLOYED FOR SOME YEARS? Fetter attributes a number of stockpile failures to an "incomplete testing program" during the development of the weapon. He writes of "haste to deploy new weapons before the Moratorium [1958–61] took effect" and suggests that "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."

It is true now, as it was in the Moratorium era, that we simply lack the resources to test weapons under all the conditions they might experience. We do the best we

^{10.} Ibid., p. 143.

^{11.} Ibid., pp. 139-140.

can and rely heavily on experience to include relevant data from similar systems. In all the problems Fetter references, weapon designers, on the basis of the information available at the time, believed that the weapons had been adequately tested. The weapons were thought to be sound. The changes made by the design laboratory in the W52 were judged to be within the range that had already been verified by nuclear testing, and it was only through nuclear testing that these changes were found to be flawed. Even experienced designers make errors, and that is one of the reasons why tests are essential.

Fetter's perspective on this is that "the technologies being deployed were new but now that the technology of nuclear weapons is more mature and well understood, problems should be less likely in the future." This skirts an issue that is of serious concern to us. We already do many fewer nuclear tests than in the 1960s. While we have learned from the problems of the 1960s, the recent problems with the W68, W80, W84, and B61 are a reminder that nuclear explosives science is still not very well understood in the 1980s.

However, Fetter misses a very important point about longevity. He is correct in noting that stockpile reliability issues have usually occurred within the first few years after deployment. In a CTB environment, we will be dealing with much longer periods of time; some repackaging of warheads will certainly be requested by the military. Repackaging of old warheads into new delivery systems with different geometries, storage conditions, and delivery-to-target sequences will introduce a second generation of problems for which we will not have nuclear certification. We agree with the statement that "problems occurred because weapon designers did not appreciate certain aspects of weapon design or the behavior of the weapon in its stockpile-to-target sequence." What concerns us is that, without nuclear certification tests, we may again fail to appreciate the important aspects. In our view, this can only increase the performance uncertainty of repackaged weapons.

Although it is easy to state that departures from experience should have been avoided during the Moratorium, circumstances mandated changes—some for safety, some for improved survivability or military effectiveness. Later in his article, Fetter suggests similar adjustments: "To solve serious reliability problems without nuclear testing, one must either remanufacture the warhead or find an acceptable substitute." Also, "if laboratory techniques or 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." Because of unique bus and interface configurations, simple RV or warhead substitutions are usually not practical (e.g., SLBMs). Some repackaging would be required. In the view of our experienced engineers and scientists, this sort of repackaging poses technical problems that must be addressed. The point here is that Fetter too recognizes there may be good reason to modify the stockpile.

^{12.} Ibid., p. 141.

^{13.} Ibid., p. 143.

^{14.} Ibid., p. 150.

To summarize, whenever problems are found, it is *ipso facto* true that "more complete" development testing might have discovered and fixed the problem. However, in the practical world of military equipment, it is impossible to test, and sometimes even to foresee, every contingency. In addition, experience has shown that the "no changes allowed" doctrine is simply not practical.

CAN REMANUFACTURE OF OLD DESIGNS SOLVE STOCKPILE RELIABILITY PROBLEMS? Fetter writes, "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. . . . 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." ¹⁵

While Fetter's argument is appealing in theory, he gives no examples of replicating decades-old technology backed up by actual experience. The issue of "replica" remanufacture is familiar to engineers in all major U.S. industries. The difficulties involved in remanufacture have been voiced by experienced people throughout government and private industry, in materials research and aerospace engineering, as well as in nuclear weapon design and fabrication. Our own experience and views regarding nuclear warhead remanufacture were confirmed by engineers responsible for the Polaris A3 missile remanufacture. Other engineers considered and subsequently rejected the proposed Saturn V remanufacture. The body of opinion regarding remanufacture can be summarized:

Exact replication is impossible to achieve. Material batches are never quite the same; some materials become unavailable; equivalent materials are never exactly equivalent; "improved" parts often fail in new and unexpected modes; vendors go out of business or stop producing some products; new health and safety regulations rule out previously used materials or processes.

Documentation cannot be made sufficiently exact to ensure replication, no matter how many resources are put into it. There can never be a perfect specification since we can never know enough to specify everything that is important. Individuals in the production plants learn how to bridge the gaps in the specifications and to make things work "regardless of what the instructions say." Documentation for older systems is especially limited. Recent weapons are extensively documented, but as noted below for the modern W68 and W84, we do not always know what to specify.

The most important aspect of any product certification is testing, because testing provides the data that make certification reliable. A nuclear test provides our only data on the nuclear performance of the whole warhead package. Testing, even with the limitations

^{15.} Ibid., p. 144.

^{16.} Miller, Brown and Alonso, Report to Congress on Stockpile Reliability.

^{17.} William J. Broad, "Hunt is on for Scattered Blueprints for Powerful Saturn Moon Rocket," New York Times, May 26, 1987.

of small numbers, is the final arbiter of competence and judgment. A major concern with nuclear weapons is that if responsible engineers were to refuse to certify a remanufactured weapon, pressures could produce individuals who would. The Challenger accident is an example of such a situation, and highlights an all-too-common tendency of human nature to allow expediency to override judgment.

Even design physicists and engineers with extensive nuclear test experience have sometimes had difficulty in correctly predicting the behavior of weapons manufactured for the stockpile. Two recent examples involve unexpected yield degradations in the production verification tests of the W68 Poseidon warhead and the W84 warhead for the ground-launched cruise missile. In fact, the W68 was a weapon remanufactured to solve a problem; besides changing the explosive, certain organic materials were no longer available and had to be changed because vendors had gone out of business. In both cases, weapon designers had predicted proper function on the basis of the production specifications, but the nuclear test results fell short of their predictions. Even in retrospect, we have been unable to explain the causes of these yield degradations on the basis of the minor changes known to exist between the development hardware and the stockpile hardware. If today, test-experienced personnel have difficulty explaining unexpected behavior in manufactured weapons, how then in the future will people without test experience be able to establish confidence in remanufactured weapons?

The role of experienced weapon designers in any remanufacturing process is crucial. Fetter incorrectly says that experienced designers are involved in the production process "only in the initial phases when it is uncertain whether the production processes can match the specifications of the designer."18 Actually, we maintain responsibility for and involvement with every weapon we design, from initial concept through production and deployment, to retirement. Weapon engineers who monitor the production of a weapon will often consult with the design physicists and get their judgments on various problems, (e.g., the acceptability of parts that are out of tolerance). Later, those who are responsible for monitoring the stockpile consult frequently with the design physicists when problems arise.

Regarding the suggestion that materials be stockpiled, we have the example of the W68 remanufacture. The organic materials that were no longer available could not have been stockpiled. Their physical characteristics required that they be formulated at the time of use. The raw materials used to make the final material had a 2-year shelf life even when refrigerated.

Fetter suggests that tightened workplace safety standards could be overcome through "additional protection for personnel . . . if national security were at stake."19 We doubt this. Once materials have been determined to be dangerous, many plants quite reasonably refuse to work on them; it is simply not practical to treat so many plastics, epoxies, and alloys with the same degree of isolation as plutonium, for example.

^{18.} Fetter, "Stockpile Confidence," p. 146.

^{19.} Ibid., p. 145.

Fetter recommends the continued remanufacture of 15–20 weapon types and the stockpiling of materials and components as a solution to the problem of remanufacture. He estimates that this would cost less than \$2 billion per year. We find this estimate to be quite unrealistic. The U.S. weapons production complex is based on aging equipment and facilities. Machines and floor space are in critically short supply; the production plants can build a few systems at a time but not the few dozen currently in stockpile. Furthermore, the unit costs Fetter infers (his note 21) are not only inaccurate (we do not have the funding to produce 4000 weapons per year) but are based on efficient production in large lots, not on slow, continuous manufacturing over 25 years.

A final note on continuous manufacturing is that people change continuously, as do gauges, glues, and welding equipment. Neither continuous manufacturing nor remanufacturing can avoid change.

COULD THE WEAPON LABS BETTER PREPARE FOR A CTB BY DESIGNING "SUPER-RELIABLE" WARHEADS AND PRE-TESTING "BACK-UP" DESIGNS? Fetter writes: "If priorities were changed so that weapon designers strove for maximum reliability instead of for the most 'bang for the buck,' it is likely that weapons could be made [more] 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."²¹

Fetter's point is not new, and has been carefully considered on a continuing basis. Our weapons already are very robust. They are as robust as we can make them within the constraints of meeting the military requirements (known officially as military characteristics or MCs) of safety, security, size, weight, yield, use of fissile material, etc. The SAAC wrote, "the Laboratories have always given great attention to stockpile endurance, and have produced designs with remarkable longevity."22 The question is, how much more robust can the designs be made? The SAAC did note that, "if the MC calling for warhead endurance were given a specific high priority, some additional robustness of design might be achieved."23 We are currently investigating the potential for more robustness in response to a request from the Senate Armed Services Committee. One thing, however, that is certain is that the concept of a completely robust or so-called "wooden bomb" is a myth. We know that adding fissile material or extra high explosive, or making the warhead heavier and larger, may provide some insurance against minor deterioration or mistakes in future evaluations, repackaging, etc. However, these measures can provide assurance against only some of these circumstances. For example, the cold IHE problem was a design flaw for which extra HE or fissile material could not have compensated. Moreover, such changes would be very expensive in terms of overall systems cost. In many

^{20.} Ibid., p. 149.

^{21.} Ibid., p. 152.

^{22.} Report of SAAC, p. 14.

^{23.} Ibid.

strategic systems, the addition of 10% more weight to the warhead has been considered and rejected. The systems penalties would be less range and therefore less operating area for missiles and submarines, or, to accomplish the same mission, 10% more ICBMs or Trident submarines. The systems cost of lower yield that Fetter would be willing to accept is either less effectiveness or more missiles. The defense and political leadership have consistently decided this trade-off on the basis of cost plus the fact that the warheads we designed have been reliable and durable.

Fetter also suggests giving up IHE or one-point nuclear safety. It is unlikely that these steps would indeed be necessary to enhance robustness. However, the consequences of giving up these peacetime safety features must be carefully analyzed and weighed against the debatable benefits of a CTB. (In a design that is one-point safe, detonation of the high explosive in an accident or terrorist attack will not produce nuclear yield.) One-point safety has been the highest priority characteristic of U.S. designs for decades. We find astonishing the suggestion that this feature be compromised in exchange for CTB robustness.

In fact, there are some practical alternatives to the unachievable robust warhead. Fetter discusses some of these, ²⁴ including the interchangeability of warheads from one missile to another, a mix of warheads on each missile type, and testing stockpiled warheads with several different kinds of HE. Fetter does not acknowledge it in his article, but these approaches are already part of U.S. design practice.

These alternatives were presented in a 1980 study and were endorsed by President Carter's Office of Science and Technology Policy (OSTP) and the National Security Council.²⁵ The OSTP Report suggested a 50% increase in nuclear testing to accommodate them. Although resources have not permitted increased testing for this purpose, we have done as much as possible on our own. Fission triggers for the cruise missile and MX warheads, for example, have been tested in several alternative configurations including different explosives. In addition, the MX and cruise missile warheads have been selected so that back-ups were already on the shelf. The land-based missiles and warheads (Mk-12, Mk-12A, Mk-21) were deliberately selected to provide the mix Fetter calls for.

Fetter claims that we could "doubtless" do more. We note that the program of nuclear tests to explore the Poseidon warhead's sensitivity and performance range would no longer be possible within the smaller number of nuclear events we now field at the Nevada Test Site. Others, not familiar with these continuing efforts to balance competing priorities and budget constraints, have accused the weapon laboratories of deliberately subverting long-standing U.S. policy by designing weapons that require continued testing to ensure their reliability. This charge was directly addressed and impartially examined by the SAAC which concluded, "this is not true."²⁶

^{24.} Fetter, "Stockpile Confidence," pp. 150–151.

^{25.} Report of the Office of Science and Technology Policy Panel on the Augmented Nuclear Test Program (ATP), The White House (May 16, 1980).

^{26.} Report of the SAAC, p. 2.

HOW IMPORTANT IS DESIGNER JUDGMENT? CAN THE WEAPONS LABS MAINTAIN IT WITHOUT NUCLEAR TESTING? Fetter addresses the subject of design experience in his discussion on remanufacturing. In our view, design experience backed up by nuclear testing is the *sine qua non* for sustaining stockpile reliability and confidence. Scientific judgment is honed by nuclear test experience.

Fetter has written about the increasing power of our computational methods. The reason these computer codes improve with time is that we are still learning about the physical processes we are trying to model. Designers perform new nuclear experiments and thus gain new knowledge, which is then incorporated into the computer models. The W58 problem is an example of this process of theory and experiment iteration: the corrosion problem could be evaluated using computer models because these models had been improved with data from nuclear tests (on a different system) for which we had better and more complete diagnostics.

Fetter asserts that the "normal stockpile surveillance program, which consists of the careful disassembly, inspection, and testing of components for many weapons, is far more effective than nuclear testing for detecting deterioration."²⁷ We agree that the surveillance program is extremely effective and catches most problems. It has frequently been noted that one-third of U.S. warhead types have required post-deployment nuclear tests to identify, evaluate, or fix stockpile problems. What should also be noted is that *all* U.S. systems have experienced real or suspected difficulties; these difficulties were resolved in *most* cases by designers who, on the basis of nuclear test experience, were able to determine that the problems could be solved without additional tests.

Fetter discusses what he believes would motivate high-quality personnel under a CTB: "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 national security." He believes that "many motivations for weapon-related work would not change significantly under a CTB." However, we view the problem from the perspective of those who have the responsibility for recruiting and developing new scientists for warhead design. The principal sources of motivation are an environment of scientific excellence and a sense of doing something important for the nation. A nuclear test ban would remove the principal scientific method from an environment of excellence and it would clearly signal that nuclear design was either unimportant or undesirable to national goals. A nuclear test ban would not only prevent us from training experienced personnel to maintain nuclear deterrence, it would also discourage high quality personnel from working in the area.

Fetter references former Director of Lawrence Livermore National Laboratory (LLNL) Roger Batzel²⁹ and writes: "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 unique disciplines of direct

^{27.} Fetter, "Stockpile Confidence," p. 142.

^{28.} Ibid., pp. 146-147.

^{29.} Ibid., p. 147.

relevance to nuclear weapons, and would retain and attract talent to the weapon laboratories."

Both Fetter's and Batzel's statements are correct. These activities can *help* motivate and maintain a staff of experienced people, and, if nuclear testing were to begin again, the new people would be further along in their understanding of the basics. However, the measures described above fall short of what is required to certify nuclear warheads. Although we have at Livermore a fairly large group of very talented ICF target designers, most of these scientists have no experience with fission primaries and strategic secondaries. In fact, because of their dedication to ICF, they are isolated from such concerns. However, such experience is crucial in order to maintain confidence when confronting the problems we would continue to have with our nuclear deterrent stockpile. These stockpile reliability issues result from manufacturing deviations, stockpile environments, and other problems unique to these weapons and outside the realm of those encountered in ICF. Furthermore, our own experience, as well as that of rocket designers and other technologists, has shown that even fully trained people need "full-up" experiments on a continuing basis to solve "new" problems of "old" systems.

We strive to maintain a continuous line of experienced designers, as senior designers pass on their knowledge to younger designers. This continuity of experience is of paramount importance. We expect that, in the event of very restrictive test limits, in only a few years we would start to lose the test-experienced people. Understandably, they would leave to work on other less restrictive, more productive projects. After a while, the people whose judgment has been honed by the realities of nuclear testing would no longer be available—they would have retired or moved on. We would then be faced with the prospect of asking scientists without nuclear testing experience to make judgments about the inevitable changes that will occur in remanufactured, repackaged, or stockpiled weapons. This is a script for disaster. If today, test-experienced personnel have difficulty explaining unexpected behaviors in the nuclear weapons they themselves have designed, how in the future will personnel without test experience be able to establish confidence in weapons designed by people long since gone?

IN THE EVENT OF LIMITS ON THE NUMBER OF NUCLEAR TESTS, HOW MANY STOCKPILE CONFIDENCE TESTS WOULD BE NEEDED TO MAINTAIN POLITICAL AND TECHNICAL CONFIDENCE IN OUR NUCLEAR STOCKPILE? Fetter writes: "Only eight of nearly 300 tests since 1970 were done to evaluate defects in stockpiled weapons. . . . One test every year or two might be enough for stockpile confidence, but not for significant modernization." ³⁰⁰

Actually, about 10% of all nuclear tests involve stockpiled hardware. In addition, it is a serious misperception to assume that only stockpile confidence tests contribute to stockpile reliability. Rather, designers draw upon the whole body of nuclear test data to ascertain the causes of nuclear weapon failures. Stockpile confidence tests

^{30.} Ibid., pp. 142, 165.

purposefully do not push the limits of our understanding, and thus they contribute less than other nuclear tests to developing designer judgment. Often, we can address stockpile reliability issues, such as the chemical degradation of the W58 Polaris warhead, without designated stockpile confidence tests; rather we can rely on test-trained people armed with data from other relevant nuclear tests. In this sense, then, every test is related to stockpile confidence.

There is another sense in which every test is a stockpile confidence test. Every test demonstrates to U.S. and Soviet leaders that the elements of the U.S. deterrent do, in fact, produce immense destructive power.

CAN LOSS OF CONFIDENCE OR WEAPON RELIABILITY IN SEVERAL U.S. STRATEGIC SYSTEMS BE MITIGATED BY STATISTICS OR BY SOVIET SELF-DOUBTS? DOES IT MAKE ANY

DIFFERENCE TO NATIONAL SECURITY? Fetter asserts that many aging problems result in reduced yield, not a complete failure.³¹ This conclusion is wrong in the majority of cases in which the reliability problem occurs with the fission trigger. For example, with the cold IHE warhead, the partial yield of the fission trigger failed to ignite the thermonuclear secondary.

Fetter assumes that there would be enough time to correct problems before most or all of the weapons are unreliable, and that the remaining weapon types would remain as a powerful deterrent. As we (and Fetter)³² have noted, there are many confidence problems in addition to those of gradual chemical deterioration, problems involving newly discovered Achilles' heels and revised stockpile-to-target sequences, storage conditions, etc. In these instances, reliability problems with an entire leg of the triad might occur suddenly. For example, new Soviet "look-down shoot-down" radars might force high-acceleration maneuvers on cruise missiles, which would, in turn, require significant warhead design changes. Even in the case of *gradual* problems, *sudden* loss of political confidence could easily result when the leadership becomes aware of the problems. The Space Shuttle is such an example: one member of the fleet proved to be unreliable at freezing ambient temperatures, but the entire fleet has been grounded because confidence was suddenly lost in the entire system.

Fetter, too, writes of the Challenger example and agrees that "the worries of engineers were overridden by program managers concerned about long delays. . . . 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." We do not find this reassuring. Nuclear reliability problems may at first be ignored by the political leaders, but then—as with the Space Shuttle—suddenly reach crisis proportions. Only a thorough nuclear testing program can avoid this dilemma.

Fetter presents the argument of those who see declining stockpile confidence as a virtue: "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

^{31.} Ibid., p. 155.

^{32.} Ibid., pp. 139-140.

^{33.} Ibid., pp. 147-148.

strike more difficult to achieve. Similarly, crisis stability would be enhanced if Soviet confidence in their stockpile eroded."34 Taken to its logical conclusion, this argument holds that we would be quite secure if both sides had no idea whether their weapons worked. We believe stability in a crisis would be less, and that the above notion is flawed. The attacker is in a much better position to exploit a large number and variety of systems than is the defender/retaliator. The first strike can allocate more than one weapon per target. Furthermore, Soviet concerns about reliability are much more likely than U.S. concerns to be ameliorated by statistics, because the Soviet deterrent is based on a much greater variety of systems.

Loss of confidence in one leg of the U.S. triad *does* have destabilizing implications. Warheads for submarine-launched ballistic missiles (SLBMs), such as the W47, W58, and W68, have experienced a number of confidence problems, so let us assume a loss of confidence in this leg of the triad during a CTB. The SLBM leg of the triad is our best insurance against a first strike in which Soviet SLBMs pin down U.S. ICBMs and destroy the alert bomber force, and then simultaneously-launched Soviet ICBMs accurately destroy U.S. silos. Clearly, loss of confidence in the one really strong leg of the U.S. triad would have disastrous security implications.

CONCLUSION

Fetter notes that "most CTB proponents want the superpowers to become less dependent on nuclear weapons for keeping the peace." Indeed, he says that "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."35

We believe this betrays a deeply rooted and serious flaw of a CTB itself as well as the CTB debate. Most of the pro-CTB consensus is based, like the nuclear freeze, on hope for an "easy way out." Responsible arms control measures recognize that strategic stability is not synonymous with static technology. Deterrence is a dynamic condition that must respond to technological developments to maintain stability.

Nuclear weapons are not unique. Like rockets, airplanes, and radar systems, nuclear weapons cannot be confidently stockpiled for several decades without testing. Like these other items, the prohibition of full-scale testing would divert scientists and engineers to other fields and confidence would further diminish. Like these other technologies, nuclear weapons must be adjusted to respond to new conditions.

We believe that more restrictive test limitations or a nuclear test ban should be considered only as part of an integrated and comprehensive approach to arms control. We must reduce the numbers of the most destabilizing weapons and the overall size of the strategic arsenals through negotiations. A restrictive test ban may be a proper eventual step in our quest for nuclear arms control and a stable peace, but it would, in our view, be an imprudent first step. Further test limitations will be consistent with increased stability and decreased tension between the U.S. and Soviet Union

^{34.} Ibid., p. 162.

^{35.} Ibid., p. 163.

only if they are instituted *after* major stabilizing reductions are made in the strategic nuclear and conventional forces of both countries.

—John D. Immele—Paul S. BrownLivermore, Calif.

The Author Replies:

In their reply to my article "Stockpile Confidence Under a Nuclear Test Ban," John Immele and Paul Brown raise many interesting points but fail to address satisfactorily the two central postulates of my article: (1) that a *combination* of nonnuclear techniques can maintain confidence with little risk of catastrophic warhead failure; and (2) that the required degree of warhead reliability depends on conceptions of nuclear strategy about which there are deep disagreements. Some maintain that highly reliable weapons are not necessary to deterrence, while even those who do demand reliable weapons should recognize that nuclear warheads are only part of a much larger and overall much less reliable military system.

Maintaining Confidence without Testing

Immele and Brown state that I advocate "that we place our confidence in . . . nonnuclear tests and computer simulations," in remanufacture, or in "repackaging."1 They then proceed to disparage each of these techniques, and to describe how each might fail. All of these techniques are fallible, of course, but the probability that none of them could resolve a given reliability problem is very low. I offered numerical estimates of this probability which Immele and Brown did not dispute. Based on past experience, there is roughly a 10% probability that a given warhead type would develop a deterioration problem, and a 25% to 50% probability that this problem could not be resolved by nonnuclear testing and computer simulation.² For the sake of argument, I assumed that remanufacture and substitution would not be effective in 10% to 20% of these cases. While I believe that the upper range of these estimates is conservative (it overestimates the probability of failure), the overall probability that all of these techniques would fail for just one of eight strategic warhead types is only about 1%. I believe that this is a small risk. Unfortunately, Immele and Brown do not offer an estimate of their own, nor do they state the degree of risk that they think would be acceptable.

First, Immele and Brown do not think that it is proper to discount design errors in the above analysis. Roughly one-third of the weapon types deployed after the Moratorium experienced problems, but only one-third of the problems were related to

^{1.} John D. Immele and Paul S. Brown, "An Exchange on Stockpile Confidence," *International Security*, Vol. 13, No. 1 (Summer 1988), p. 199.

Security, Vol. 13, No. 1 (Summer 1988), p. 199. 2. Steve Fetter, "Stockpile Confidence Under a Nuclear Test Ban," International Security, Vol. 12, No. 3 (Winter 1987/88), p. 164.

deterioration; the remainder were due to design errors.³ It is a fact, however, that all known design errors have been detected and resolved within four years of the date that the warhead type was first produced; indeed, except for three problems in the early 1960s, all were fixed within two years of first production.4 The newest strategic warhead—the W87 for the MX missile—is now two years old; other strategic warheads have been deployed for five to twenty-five years.⁵ If we are prudent enough (as we were not during the 1958-61 Moratorium) to rely only on those warheads that have been stockpiled for several years when a test ban begins, then experience shows that design errors can be avoided.

Second, Immele and Brown claim that, because modern thermonuclear weapons rely on "boosting," which is a detailed process that cannot be adequately modeled using computer simulations or nonnuclear testing, many stockpile confidence problems are unlikely to be resolved using only those techniques. The 25% to 50% probability given above that these techniques would fail to resolve a problem was meant to indicate their limitations. The important point is that, in the past, several problems have been resolved without nuclear testing (or remanufacture or substitution). It is not "naive" to believe that nonnuclear solutions will continue to be valuable in this regard. Nowhere did I suggest, as they charge, that "computer simulations [alone] can replace nuclear testing";6 if they could, a test ban obviously would be worthless as an arms control measure.

Third, Immele and Brown state that while remanufacture is "appealing in theory," I give no examples "backed up by actual experience." In fact, I clearly state that one U.S. warhead (the identity of which is classified) was successfully remanufactured after the production line was shut down for three years. No nuclear tests were done.8 Immele and Brown imply that remanufacture is impossible, but a study of the technical feasibility of remanufacture has never been done. Exact replication is impossible, of course, but could it be close enough? How difficult and costly would it be to keep assembly lines open? To persuade vendors that it is their national duty to continue to produce certain materials? To subsidize these vendors, or to produce the materials in government laboratories? To certify alternate materials for key components? In testimony before Congress, both Admiral Sylvester R. Foley (former Assistant Secretary of Energy for Defense Programs) and Roger Batzel (former Director of Lawrence Livermore National Laboratory) have stated that, given enough time and money, remanufacture could be achieved.9

^{3.} Ibid., p. 162 n. 42.

^{4.} Ibid., Table 1, p. 137.

^{5.} Thomas B. Cochran, William M. Arkin, Robert S. Norris, and Milton M. Hoenig, Nuclear Weapons Databook, Vol. II, U.S. Nuclear Warhead Production (Cambridge, Mass.: Ballinger, 1987), pp. 10–11.

^{6.} Immele and Brown, "Commentary," p. 200.

^{7.} Ibid., p. 202.

^{8.} Fetter, "Stockpile Confidence," p. 143.

^{9.} 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, 19 February 1986. Roger Batzel, responses to questions before the Armed Services Committee, U.S. Senate, April 1986.

Fourth, Immele and Brown say almost nothing about the feasibility of substitution, although at one point they assert that the "repackaging of old warheads into new delivery systems" may introduce more problems than it solves. ¹⁰ In fact, I never mentioned "repackaging" as a solution to stockpile confidence problems; instead, I referred to substituting an entire bomb or warhead (in its original reentry vehicle) for the problem bomb or warhead. For example, the two Minuteman III warheads and the MX warhead are perfect substitutes for one another. In addition, the B61 and B83 bombs, the two cruise missile warheads, and the Trident I and Trident II warheads could substitute for each other. No repackaging is necessary, and the military utility of these systems is affected little by these substitutions. Of course, the substitute might fail as well, but the probability that two warheads of a given class would fail is very much smaller than the probability that just one would fail.

At the outset of my discussion, I defined stockpile confidence to mean "a reasonable assurance that the stockpile of nuclear weapons already deployed at the time a CTBT [comprehensive test ban treaty] is negotiated would continue to perform, with existing delivery systems, with the degree of reliability necessary for deterrence." Repackaging is not necessary in order to use existing warheads on different existing delivery systems when the stockpile-to-target sequences (STSs) of these systems are similar. In fact, with minor changes the Minuteman/MX warheads could be exchanged with Trident warheads. Repackaging existing warheads for use with new delivery vehicles is modernization, not stockpile maintenance. High confidence should nevertheless be possible if the STS of the new delivery vehicle is not outside the range already verified by nuclear testing. In contrast, high confidence in repackaging would not be possible when the STS is very different from that of existing delivery vehicles; this is one reason proponents hope a CTB would prevent the development of new weapons that they believe are destabilizing (e.g., earth-penetrating warheads).

Fifth, Immele and Brown belittle the idea of designing super-reliable warheads on the grounds that they would be expensive and less effective, because even super-reliable warheads would not be perfect, and because current warheads are already very reliable. I never stated that current warheads were unreliable or that a program to design super-reliable warheads should be undertaken, nor did I recommend, if such a program was completed, that super-reliable warheads be stockpiled. Super-reliable warheads are an option for those who feel, for whatever reasons, that non-nuclear techniques, remanufacturing, and substitution would be inadequate to resolve most stockpile confidence problems. Such a program need not be excessively expensive since only three strategic weapons need to be developed: one for ballistic missiles, one for bombs, and one for air-launched missiles. The weapons need not be deployed in large numbers; the existence of super-reliable designs would simply be an insurance policy against catastrophic failure in the regular stockpile. Thus, any decrease in the safety or effectiveness of the stockpile associated with the introduction of super-

^{10.} Immele and Brown, "Commentary," p. 201.

^{11.} Fetter, "Stockpile Confidence," p. 133 (emphasis added).

^{12.} Personal communication, Steven Peglow, Lawrence Livermore National Laboratory, 2 May 1988.

reliable warheads need only be accepted in the unlikely event that other measures short of nuclear testing could not solve the problem.

Super-reliable weapons need not be markedly inferior to regular weapons. Immele and Brown recoil at the notion of "giving up" insensitive high explosive (IHE) or onepoint safety. In fact, I only suggested relaxing the (arbitrary) one-point safety criterion, not giving up the requirement for a reasonable degree of nuclear safety altogether. And why should IHE be used in super-reliable weapons if they are unlikely ever to be deployed extensively, especially when use of IHE has resulted in its own problems? Conversations with experienced weapon designers have convinced me that weapons far less vulnerable to certain failure modes can be made without substantial increases in weight or volume and without significant sacrifices in safety or security. Superreliable weapons would probably be more expensive than normal weapons, but this does not appear to be a decisive factor against such a program.

Finally, Immele and Brown do not address themselves to the cost or feasibility of programs to certify alternate materials for key warhead components. If, for example, high explosives are known to deteriorate or to be difficult to replicate, then why not certify two types of high explosive for each key warhead design?

STOCKPILE CONFIDENCE AND NUCLEAR STRATEGY

Immele and Brown state that they will address "whether weapon confidence can keep pace with the requirements of U.S. nuclear deterrence,"13 but they do not discuss just what they mean by "nuclear deterrence" or exactly what weapon capabilities they think are required for deterrence. A multitude of distinct strategies falls under the broad category of "nuclear deterrence." From their discussion, however, it is clear that Immele and Brown subscribe to what may be loosely described as the "warfighting" or "war-winning" school of deterrence theory. In general, those who favor war-fighting deterrence require high confidence in the reliability of weapons because they feel that a counterforce attack could materially limit damage to the United States by destroying the capability of the Soviet Union to wage war. In addition, they feel that maintaining war-fighting capabilities improves deterrence by frightening the Soviets so that they would be more reluctant to engage in actions that might provoke hostilities in the first place.

Proponents of war-fighting deterrence are natural enemies of a test ban because a CTB would hinder weapon modernization, not because it would significantly degrade stockpile reliability. Nuclear weapons represent the strongest link in a weak chain of nuclear war-fighting capabilities. As I showed in my article, the uncertainty in the performance of nuclear warheads is a tiny ripple in the turbulent sea of uncertainties surrounding the performance of the whole military system.¹⁴ This is especially true when one contemplates the execution of dubious war-fighting missions, such as the preemptive destruction of an entire intercontinental ballistic missile force. The ability to carry out sophisticated war-fighting missions with nuclear forces depends, for

^{13.} Immele and Brown, "Commentary," p. 197.

^{14.} Fetter, "Stockpile Confidence," pp. 153–160.

example, on the survival of satellites whose vulnerability is unknown. It is only natural to want the highest possible confidence in our military systems, but concern about warhead reliability is like worrying about the reliability of the on-off switch in an untested supercomputer.

Immele and Brown state that "The consequences of being wrong [about stockpile confidence during a CTB] are monumental."15 Later, they assert that a "loss of confidence in the one really strong leg of the U.S. triad [the submarine-based force] would have disastrous security implications."16 What are these "monumental" and "disastrous" implications? Are they suggesting, as the logic of their argument indicates, that a loss of confidence in a single warhead could lead to nuclear blackmail or, worse, to nuclear war? If so, then Immele and Brown have lost touch with political reality. In a recent study, Richard Betts found that the propensity for nuclear blackmail was unrelated to the nuclear balance, and that the success of nuclear threats was only weakly correlated with the strategic balance.¹⁷ Fortunately, there is no plausible scenario by which a loss of confidence would lead to nuclear war. Even if U.S. confidence in a particular warhead type had waned, what rational Soviet leader would be foolish enough to test the actual reliability of these warheads? Indeed, what Soviet leader would then want to test the actual reliability of his own weapons? Even if Soviet leaders were somehow convinced that their own weapons were reliable and that our submarine-based weapons were unreliable, what Soviet leader would assume that our land-based missiles would wait patiently in their silos while under attack?

One should not forget that many thoughtful people do not believe that maintaining the capability to fight and win a nuclear war is the best way to prevent war. They see increased counterforce capabilities—which continued nuclear testing helps make possible—as dangerously destabilizing because they can increase the premium for striking first during a crisis and because they can lead to an arms competition. More moderate deterrent strategies that seek to avoid these instabilities while preserving a "credible" threat for nuclear use, such as "soft counterforce" (targeting troop concentrations or certain military installations), do not require extremely reliable weapons. Moreover, considerations about stockpile reliability are completely irrelevant for those who feel that even limited counterforce options are unwise. Advocates of "minimum deterrence," for example, claim that the possibility that a few dozen cities would be destroyed is enough to deter an adversary. It is not surprising that Immele and Brown fail to address the degree of reliability necessary for strategies that require much more modest nuclear capabilities than those that we currently possess.

SUMMARY

I know of no person whose opinion about the desirability of a test ban hinges on the question of stockpile reliability. Now, as in the past, such opinions are based on the desire to develop, or to halt the development of, new nuclear warhead types. These

^{15.} Immele and Brown, "Commentary," p. 197.

^{16.} Ibid., p. 209.

^{17.} Richard K. Betts, Nuclear Blackmail and Nuclear Balance (Washington, D.C.: Brookings, 1987).

Correspondence | 215

desires, in turn, are based on conceptions of nuclear strategy and on broad political ideas about how best to manage the competition between the superpowers and to limit the spread of nuclear weapons to additional nations. It is time to stop arguing about technical details and to start an honest dialogue about these more fundamental issues.

—Steve Fetter Boston, Mass.