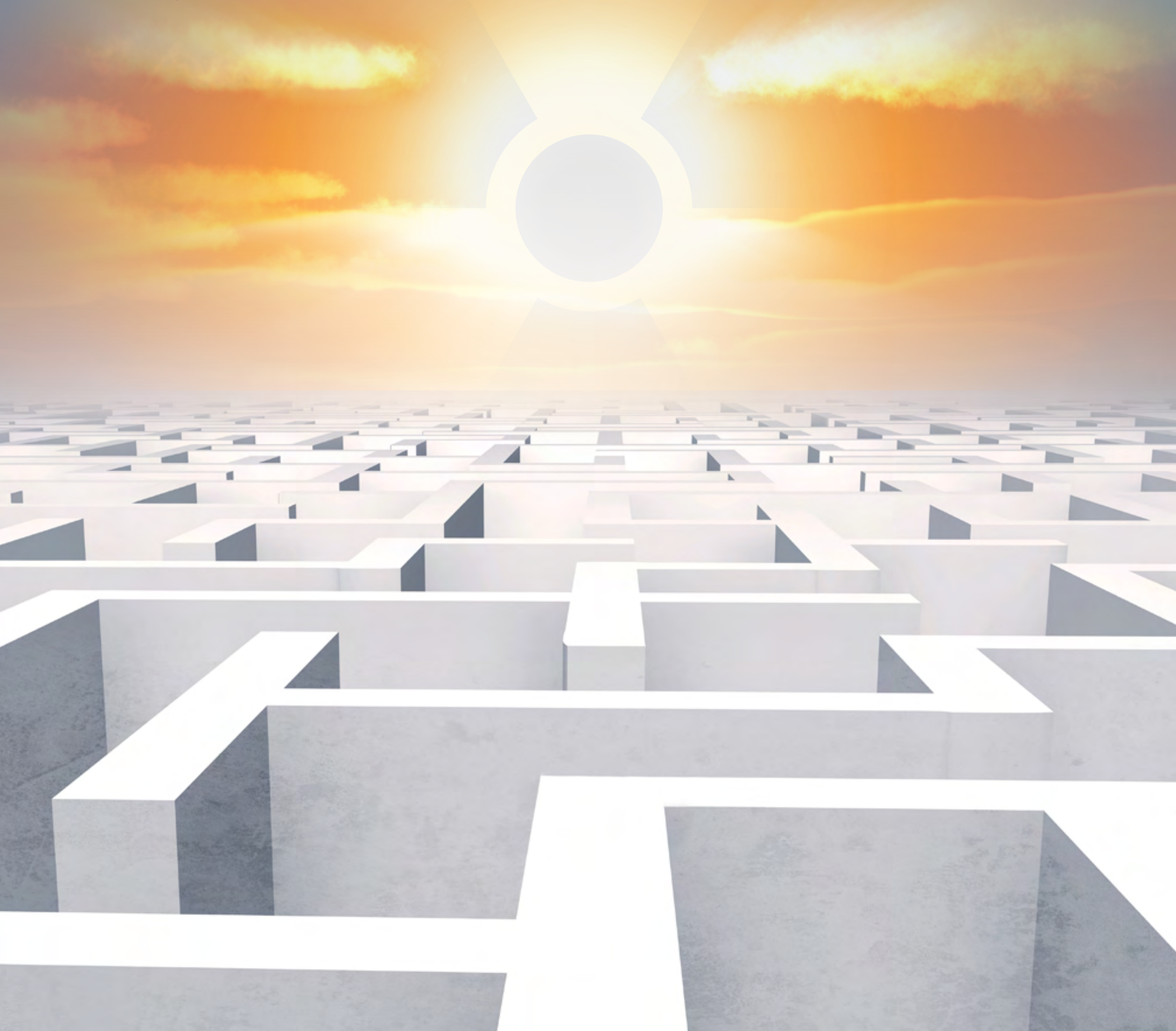
Technical Issues in the Comprehensive Nuclear Test Ban Treaty (CTBT) Ratification Debate: A 20 YEAR RETROSPECTIVE

Anna Péczeli and Bruce Goodwin

Center for Global Security Research
Lawrence Livermore National Laboratory

September 2020



TECHNICAL ISSUES IN THE COMPREHENSIVE NUCLEAR TEST BAN TREATY (CTBT)

RATIFICATION DEBATE

A 20 YEAR RETROSPECTIVE

ANNA PÉCZELI & BRUCE GOODWIN

Center for Global Security Research
Lawrence Livermore National Laboratory
September 2020

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory in part under Contract W-7405-Eng-48 and in part under Contract DE-AC52-07NA27344. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC.
ISBN-978-1-952565-05-2 LCCN-2020917894 LLNL-TR-813677 TID-60015

Table of Contents

About the Authors

Bruce Goodwin is a senior fellow at the Center for Global Security Research at Lawrence Livermore National Laboratory (LLNL). Prior to joining CGSR, he was LLNL's associate director-at-large for national security policy and research and from 2001 to 2013 was the principal associate director for LLNL's nuclear weapons program. In this latter role, Dr. Goodwin led the process to certify LLNL's nuclear weapons and was responsible for establishing priorities, developing strategies, and designing and maintaining LLNL's nuclear weapons. From 1981 to 2001, he worked as a nuclear weapons designer at Los Alamos and at Lawrence Livermore, and authored five nuclear weapons tests. For this work he received the Department of Energy E.O. Lawrence Award for innovative weapons science. Goodwin received his bachelor's degree in physics from City College of New York and his master's and doctorate in aerospace engineering from the University of Illinois at Urbana Champaign.

Anna Péczeli is a postdoctoral research fellow at the Center for Global Security Research (CGSR). She is also an affiliate at the Center for International Security and Cooperation (CISAC) at Stanford University and an adjunct fellow at the Centre for Strategic and Defence Studies (CSDS) at the National University of Public Service in Budapest, Hungary. Prior to joining CGSR, Dr. Péczeli was a visiting postdoctoral research scholar at Stanford University's Europe Center (2017-2018) and a Stanton nuclear security fellow at CISAC (2016-2017). In Hungary, she was a senior research fellow at CSDS, an assistant lecturer at Corvinus University of Budapest, and an adjunct fellow at the Hungarian Institute of International Affairs. During her Ph.D. studies, Péczeli held a visiting research fellowship at the Peace Research Institute Frankfurt and was a visiting Fulbright fellow at the Nuclear Information Project of the Federation of American Scientists (FAS) in Washington, DC. She earned a Ph.D. in international relations from Corvinus University of Budapest.

Brad Roberts is director of the Center for Global Security Research at Lawrence Livermore National Laboratory, a position he assumed in 2015. From 2009 to 2013, he served as deputy assistant secretary of defense for nuclear and missile defense policy. In this capacity, he served as policy co-director of the Obama administration's Nuclear Posture Review and Missile Defense Review. From 2013 through 2014, Dr. Roberts was a consulting professor at Stanford University and William Perry Fellow at the university's Center for International Security and Cooperation. While there, he authored *The Case for U.S. Nuclear Weapons in the 21st Century* (Stanford University

Press, 2015), which was subsequently recognized with a Choice Award for outstanding academic publication of the year.

From 1995 to 2009, he was a member of the research staff at the Institute for Defense Analyses in Alexandria, Virginia. From 1983 to 1995, he was a research fellow at the Center for Strategic and International Studies, where he also served as editor of the *Washington Quarterly* from 1986 to 1995. He holds a B.A. from Stanford University, an M.Sc. from the London School of Economics and Political Science, and a Ph.D. from Erasmus University, Rotterdam.

Preface

Brad Roberts

In its 2009 report on U.S. nuclear strategy, the bipartisan Congressional Commission on the Strategic Posture of the United States agreed to a total of 95 findings and recommendations. Only one hard problem would not give way to the pressures of former Secretaries of Defense William Perry and James Schlesinger, the Commission's formidable chair and vice chair. That stubborn issue was, of course, the Comprehensive Nuclear Test Ban Treaty (CTBT). Debate over the CTBT has been as prolonged and rancorous as any debate in U.S. national security policy. Resolution of that debate appears nowhere in sight.

But context and circumstance change. Thus, it is useful periodically to review and update the terms of debate. The CTBT debate has been infused with a complex mix of policy, political, strategic, military, and technical perspectives. As a research center sitting at the intersection of the technical and policy worlds, the Center for Global Security Research saw an opportunity. The idea for this study took shape as we followed the renewed debate about the CTBT on the occasion of the 20th anniversary of the Senate's rejection of consent to ratification. While we've been at work, new factors have driven the CTBT back into debate in the policy community—developments that post-date our decision to prepare this paper.

Our focus is narrowly technical. After reviewing the CTBT ratification debate, we chose five technical issues as especially significant in the debate and set out to explore what we have learned in the interim. This paper reviews the issues and what we learned. It does not review the policy or political issues in the ratification debate. Nor do we offer recommendations for the future of the CTBT.

I am grateful to the co-authors for tackling such a complex and politicized topic. They have done a commendable job of analyzing the issues in a clear and dispassionate way. The views expressed here are their personal views and should not be attributed to Lawrence Livermore National Laboratory, the National Nuclear Security Administration, or any other entity.

Introduction

Two decades after the U.S. Senate declined in 1999 to give its consent to ratification of the Comprehensive Nuclear Test Ban Treaty (CTBT), the treaty remains in legal and political limbo. Legally, the treaty cannot enter into force without U.S. ratification. However, the Senate's rejection does not strip the treaty of its force as an element of customary international law or alter the will of other signatories to see it enter into force. Politically, the debate over the virtues and flaws of the treaty remains as strong as ever, with some CTBT advocates pushing energetically for U.S. re-review while some opponents seek ways to "un-sign" the treaty. Meanwhile, despite the Senate's rejection of the treaty, the United States acts in a manner consistent with its main obligations by maintaining a moratorium on explosive nuclear testing and providing financial support (as well as technical expertise) to the CTBT Organization (CTBTO)—in particular the CTBTO's International Monitoring System (IMS).

The CTBT may well remain in such limbo for a long time to come, unless some catalytic event comes along to re-shape national perceptions and priorities, one way or the other. The treaty's future, whatever it might be, will be determined in part by how much perspectives might have changed in light of interim developments. Our purpose is to shed light on some of those developments and to assess their relevance.

The 1999 ratification debate addressed a large number of issues. Rejection was driven by various policy and technical judgments. The purpose of this paper is to re-examine those technical judgments in light of what has been learned over the following two decades. On some technical matters, there was explicit recognition of extant uncertainties, as well as a conviction and hope that time would dispel them. To be doubly clear, our purpose is not to recommend a way forward on CTBT ratification. The policy judgments of 1999 and 2020 are beyond the scope of this paper.

This technical re-examination proceeds as follows. It begins with a short background section to introduce the key technical matters in discussion in the U.S. ratification process. Then it moves systematically through each of those topics with a review of relevant background information, a discussion of key points in debate, an analysis of subsequent experience, and a current assessment. The historical analysis draws on many sources including the primary documents of the Congressional debate, the official article-by-article interpretation by the Department of State, and the statements of key players in the process. We also relied on key technical reports and publications by subject matter experts that have addressed in more detail many of the issues in our focus. These sources include the 2002 and 2012 reports of the National Academy of Sciences[1] as well as publications by senior technical staff of Lawrence Livermore National Laboratory (LLNL) who were actively involved in the ratification debate, including then-Laboratory Director Bruce Tarter.[2] We also conducted several interviews with government officials from the 1990s, past laboratory directors, and other members of the scientific community. But, except where individuals are specifically cited by name, the judgments offered here are our own and should not be attributed to any other individual or to LLNL as such.

1 National Academy of Sciences, *Technical Issues Related to the Comprehensive Nuclear Test Ban Treaty* (Washington, DC: The National Academies Press, 2002). https://doi.org/10.17226/10471. Accessed September 15, 2020.
National Research Council, *The Comprehensive Nuclear Test Ban Treaty: Technical Issues for the United States* (Washington, DC: The National Academies Press, 2012). https://doi.org/10.17226/12849. Accessed September 15, 2020.

2 Tarter, Bruce C., *The American Lab: An Insider's History of the Lawrence Livermore National Laboratory* (Baltimore, MD: Johns Hopkins University Press, 2018); and Brown, Paul, *The Comprehensive Test Ban Treaty: Lawrence Livermore National Laboratory's Impact on U.S. Nuclear Policy from 1958 to 2000* (Livermore, CA: Lawrence Livermore National Laboratory, 2019). https://cgsr.llnl.gov/content/assets/docs/Brown-CTBTbook.pdf. Accessed September 15, 2020.

Background

Initiatives to stop or limit nuclear testing date back to the dawn of the nuclear age. Negotiations on a ban commenced in 1958, in conjunction with the first test moratorium jointly implemented by the United States, Soviet Union, and United Kingdom. The first limitation was agreed to in 1963; the Limited Test Ban Treaty (LTBT) prohibited the testing of nuclear weapons in outer space, underwater, or in the atmosphere. In 1974 the Threshold Test Ban Treaty (TTBT) prohibited tests with a yield of more than 150 kilotons. The 1976 Peaceful Nuclear Explosions Treaty (PNET) governs nuclear explosions that are carried out outside test sites, as established under the TTBT; with it, the United States and the Soviet Union agreed to not carry out explosions exceeding 150 kilotons and group explosions unless the yield of individual explosions in the group could be verified.

After the conclusion of the TTBT and the PNET, the United States, United Kingdom, and Soviet Union in 1977 began negotiations aimed at concluding a comprehensive test ban (CTB). At the time, there were competing interpretations of the purpose and objective of such a ban.[3] Some saw it primarily as a tool of disarmament and an "adjunct to Article 6" of the Nuclear Non-proliferation Treaty (NPT). This line of thinking was based in the belief that if nuclear powers are not allowed to test their weapons, they will not be able to maintain and modernize their arsenals, and that this would eventually lead to nuclear disarmament. Others saw it primarily as a tool of non-proliferation, both horizontal (proliferation among nations) and vertical (proliferation within nations as they pursue qualitative and quantitative improvements to their forces). A third way of thinking emphasized the value of a test ban as a tool of deterrence: the CTB would only delegitimize nuclear testing, but not the possession of nuclear weapons. Thus, it might cap the qualitative arms race and reduce incentives for any quantitative arms race, all while permitting the sustainment of deterrence.

Two decades later, in 1996, the UN General Assembly approved the CTBT. The preamble of the treaty[4] underlines that the CTBT is an important nuclear non-proliferation and disarmament tool. Its main provisions are the prohibition of carrying out any nuclear explosion at any place under the jurisdiction or control of state

---

3 Interview with Amb. Ron Lehman, director of the Arms Control and Disarmament Agency (1989-1993) (January 16, 2020).
Johnson, Rebecca. *Unfinished Business: The Negotiation of the CTBT and the End of Nuclear Testing* (United Nations Institute for Disarmament Research, Geneva: UNIDIR Press, 2009). https://www.unidir.org/files/publications/pdfs/unfinished-business-the-negotiation-of-the-ctbt-and-the-end-of-nuclear-testing-346.pdf. Accessed September 15, 2020.
Hansen, Keith A., *The Comprehensive Nuclear Test Ban Treaty – An Insider's Perspective* (Stanford, CA: Stanford University Press, 2006).

4 United Nations, *Comprehensive Nuclear-Test-Ban Treaty* (CTBT) (September 24, 1996).
https://www.ctbto.org/fileadmin/content/treaty/treatytext.tt.html. Accessed September 15, 2020.

parties, and the prohibition of any encouragement of or participation in the carrying out of any nuclear explosion.

As the treaty opened for signature on September 24, 1996, the United States was first in line. In 1997, President Clinton sent the CTBT to the Senate for its advice and consent to ratification. The Senate Committee on Armed Services held hearings[5] on October 6-7, 1999; the Senate Committee on Foreign Relations held hearings[6] on October 7, 1999; and the Senate held an Executive Session[7] on October 8, 1999. On October 13, 1999, the Senate rejected consent to ratification with a closely divided vote. Forty-eight senators voted in favor and 51 voted against, with those in favor falling well short of the 67 votes required for consent to ratification. As of 2020, the treaty has 184 signatories and 168 ratifications.[8] Of the 44 countries identified in Annex 2 of the Treaty that must submit instruments of ratification for the treaty to enter into force, 36 have signed and ratified, and an additional five states have signed but not ratified (one of them is the United States).[9]

The 1999 ratification debate illuminated strongly conflicting perspectives on the utility and effectiveness of the treaty in stopping foreign nuclear weapon developments and on its impact on the ability of the United States to maintain credible nuclear deterrent over the long term. That debate has been well summarized by others.[10]

5 United States Senate, "National Security Implications of the Comprehensive Test Ban Treaty" (October 6-7, 1999). Hearing before the Committee on Armed Services, S. Hrg. 106-490.

6 United States Senate, "Final Review of the Comprehensive Nuclear Test Ban Treaty" (October 7, 1999). Hearing before the Committee on Foreign Relations, S. Hrg. 106-262. https://www.govinfo.gov/content/pkg/CHRG-106shrg61364/html/CHRG-106shrg61364.htm. Accessed September 15, 2020.

7 United States Senate, "Comprehensive Nuclear Test-Ban Treaty," Senate Executive Session (October 8, 1999). https://www.congress.gov/congressional-record/1999/10/08/senate-section/article/S12257-6. Accessed September 15, 2020.

8 Comprehensive Nuclear-Test-Ban Treaty Organization, "Status of Signature and Ratification" (April 2020). https://www.ctbto.org/the-treaty/status-of-signature-and-ratification/. Accessed September 15, 2020.

9 As of Summer 2020, of the 44 countries required for entry into force, the following have not submitted instruments of ratification: China, the Republic of Korea, Egypt, India, Iran, Israel, Pakistan, and the United States of America. See Appendix 1 for a chronology of key developments bearing on the CTBT and its ratification in the United States.

10 Clinton, William J., "Remarks on the Comprehensive Nuclear-Test-Ban Treaty and an Exchange with Reporters," The White House, Office of the Press Secretary (July 20, 1999). https://www.presidency.ucsb.edu/documents/remarks-the-comprehensive-nuclear-test-ban-treaty-and-exchange-with-reporters. Accessed September 15, 2020.
Roehl, Jayson, "The United States Senate and the Politics of Ratifying the Comprehensive Nuclear-Test-Ban Treaty," *Comparative Strategy* 28, no. 4 (2009), p303-316.
Kimball, Daryl, "What Went Wrong: Repairing Damage to the CTBT," *Arms Control Today* 29, no. 8 (December 1999). https://www.armscontrol.org/act/1999-12/features/what-went-wrong-repairing-damage-ctbt. Accessed September 15, 2020.
Deibel, Terry L., "The Death of a Treaty," *Foreign Affairs* (September/October 2002). https://www.foreignaffairs.com/articles/united-states/2002-09-01/death-treaty. Accessed September 15, 2020.
Nikitin, Mary Beth D., *Comprehensive Nuclear-Test-Ban Treaty: Background and Current Developments* (Washington, DC: Congressional Research Service, RL33548, September 1, 2016). https://crsreports.congress.gov/product/pdf/RL/RL33548. Accessed September 15, 2020.

Technical issues featured prominently in the ratification debate. Anticipating that debate, the Clinton administration committed to a set of safeguards as it prepared the ratification process. Those safeguards included the following provisions:[11]

1. the conduct of a science-based Stockpile Stewardship Program and a broad range of effective and continuing experimental programs;
2. the maintenance of modern nuclear laboratory facilities and programs;
3. the maintenance of the basic capability to resume nuclear test activities;
4. the continuation of a comprehensive research and development program to improve treaty monitoring capabilities and operations;
5. the continuing development of intelligence gathering and analytical capabilities and operations to ensure accurate and comprehensive information on worldwide nuclear arsenals;
6. the President would be ready to withdraw from the CTBT under the “supreme national interests” clause if a high level of confidence in the safety or reliability of the U.S. nuclear weapons stockpile could no longer be certified.

The Senate Minority Whip, Jon Kyl, played the leading role opposite the administration and in the course of the ratification debate raised numerous questions about its technical merits.[12] Among the many technical issues then in discussion, the following stand out as especially prominent and consequential for the balance of Senate opinion:

1. Would stockpile stewardship alone provide the needed confidence in the reliability and safety of the existing arsenal?
2. Would it provide the needed confidence in any modernized or new weapon that might be developed without nuclear testing?
3. Is it possible to maintain the needed expertise to test in the absence of testing?
4. Can the combined monitoring capabilities of the United States and IMS provide the information necessary to verify compliance by all state parties, including detection of low-yield and decoupled tests?

11 Clinton, William J., “Message to the Senate Transmitting the Comprehensive Nuclear Test-Ban Treaty with Documentation,” The White House, Office of the Press Secretary (September 22, 1997). https://www.presidency.ucsb.edu/documents/message-the-senate-transmitting-the-comprehensive-nuclear-test-ban-treaty-with. Accessed September 15, 2020.

12 In September 1997, Senator Kyl sent a letter to LANL and LLNL laboratory directors, containing 21 questions about the stockpile stewardship program and the CTBT. See the list of questions in Appendix 1. The responses from LANL Director Siegfried S. Hecker and LLNL Director Bruce Tarter are part of the Congressional record of the October 27, 1997 Senate hearing on the safety and reliability of the U.S. nuclear deterrent.
United States Senate, “Safety and Reliability of the U.S. Nuclear Deterrent” (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267. https://www.govinfo.gov/content/pkg/CHRG-105shrg44720/html/CHRG-105shrg44720.htm. Accessed September 15, 2020.

5. Has the United States put itself at a disadvantage with its commitment to the zero-yield standard?

On each of these questions, the views of the technical expert community were sought.[13] Members of Congress, of course, formed their own judgments and voted accordingly.

Twenty years later, proponents and opponents remain deeply divided.[14] But the passage of time has brought much interim experience that can inform our collective thinking about the nuclear testing moratorium and the merits and deficiencies of the CTBT. Indeed, it is possible to revisit the questions noted above with an improved understanding born of interim experience. The remainder of this paper addresses each of these questions in turn.

13 Ibid.

14 Some opponents are exploring what can be done to remove any further possibility of a future re-review of the CTBT by the United States Senate. See Appendix 3 for a discussion of the effort to “un-sign” the treaty.

Ensuring Confidence in the Safety and Reliability of Existing Nuclear Weapons

The Senate wanted to understand the impact of a permanent ban on testing on the ability of the United States to retain confidence in the safety and reliability of its existing nuclear arsenal. As Senator Kyl put it in his 1997 letter to the laboratory directors, "Will confidence in the safety and reliability of U.S. nuclear weapons decline in the absence of testing?"[15]

For decades, the United States had relied on nuclear testing for that purpose. As the Department of Energy noted in its official history of the test program, the goals of nuclear testing were: "to prove that the weapon would work as designed…or to advance nuclear weapon design, or to determine weapons effects, or to verify weapon safety."[16] Towards these ends, the United States conducted altogether 1,030 nuclear tests, beginning on July 16, 1945 and concluding on September 23, 1992.[17] These tests were integral to the approach the United States had taken from the start to the development and maintenance of the nuclear deterrent—what might be called the "cut-and-try" methodology. In this approach, a device was designed and built using the best expertise of designers and engineers. Then it was tried in a nuclear test to determine if and how it would perform. Weapons must work under varied deployment conditions and over an extended lifetime. During that lifetime, radioactive materials, such as tritium, decay. Hence, multiple tests under these various conditions were required for each weapon type. This continuous cycle came to a stop in 1992 when Congress mandated a pause in nuclear testing, and then President William J. Clinton decided to pursue a zero-yield CTBT to make it a permanent, legally binding obligation.

In its ratification review, the Senate wanted to understand whether the Stockpile Stewardship Program (SSP), then being pursued as an alternative means to ensure confidence in safety and reliability, would prove fit for purpose. Again, in Senator Kyl's words, "What are the specific measures by which you will know whether [SSP] will have succeeded or failed?"[18]

15 Kyl, Jon, "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

16 United States Department of Energy, *United States Nuclear Tests – July 1945 through September 1992* (DOE/NV--209-REV 16, September 2015), pvii. https://www.nnss.gov/docs/docs_LibraryPublications/DOE_NV-209_Rev16.pdf. Accessed September 15, 2020.

17 This data does not include the 24 tests that the United States held jointly with the United Kingdom. It also does not include the two nuclear weapons that the United States exploded over Japan in 1945.

18 Kyl, Jon, "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

The Clinton administration crafted and promoted the SSP to sustain the deterrent in the absence of nuclear testing. SSP promised to substitute advanced computational and diagnostic capabilities for underground tests. As Los Alamos National Laboratory Director Siegfried S. Hecker explained: "The premise of SBSS [science-based stockpile stewardship] is that we can offset the loss of confidence in the safety and reliability of existing weapons without nuclear testing by demonstrating improved understanding of the underlying science and technology."[19]

The ensuing debate reflected some degree of uncertainty within the technical community about the prospects for the success of stockpile stewardship. At the time, stockpile stewardship was a grand but very speculative endeavor. The two nuclear weapons design laboratories (Lawrence Livermore National Laboratory and Los Alamos National Laboratory, LLNL and LANL) understood what was needed to experimentally and computationally detail the operation of a nuclear weapon, so there was the potential for success. But this was seen as a "grand challenge." Success was not guaranteed either technically or financially.

Uncertainty derived in part from the challenge of assembling the needed infrastructure. As LLNL Director, Tarter noted in his letter to Senator Kyl, "the major challenge lies ahead, more powerful computers, advanced experimental facilities, modern manufacturing facilities and enhanced surveillance capabilities are required to deal with inevitable aging problems in the stockpile and to demonstrate unambiguously our level of expertise to make judgments about the stockpile."[20] The success of SSP depended on multiple new facilities to experimentally probe the details of nuclear function—each on an enormous scale. The main components of the envisioned architecture were facilities to:

- Understand fusion ignition [this became the National Ignition Facility (NIF) at LLNL]
- Understand pre-ignition material behavior [this became the Dual Axis Radiographic Hydrodynamic Test (DARHT) Facility at LANL and the Contained Firing Facility (CFF) at LLNL]
- Understand the behavior of plutonium under shocked conditions [this became, for example, the Subcritical Experimental Program and the Joint Actinide Shock Physics Experimental Research Facility (JASPER) at the Nevada National Security Site]
- Run advanced simulations with improved data (this became the Advanced Simulation and Computing Initiative—ASCI)

19 Hecker, Siegfried S., "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

20 Tarter, Bruce C., "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

Individually, these were each expensive projects. It was confirmed at the 1997 Senate hearing on the SSP that the Clinton administration committed to fund stockpile stewardship at about $4.5 billion in FY 1999 and to use that as a baseline for future funding.[21] Taken together, the total national investment over two decades would come to tens of billions of dollars.[22]

Next, there was the question of whether it was even technically feasible for American industry to develop the enormously large and fast computers needed to run the vast numbers of exquisitely resolved weapon simulations that would be developed and validated with data from the expensive experimental facilities. Could the nation's computer industry succeed in developing machines that were more than a million times faster than the machines of 1995? Again, would the budgets to fund that program, i.e. the ASCI, be sustainable over the decades of development?

In making its case to test-ban skeptics in the U.S. Senate, the Clinton administration emphasized the investments it planned for SSP. It also emphasized that many activities were still allowed under the CTBT that were part of the plan to maintain the confidence in the U.S. stockpile. On top of these arguments, the White House also agreed to an unprecedented safeguards package. The administration decided that the success of the SSP would be reviewed annually, and if any of the three laboratory directors said that they cannot maintain the nuclear stockpile without nuclear testing, the Administration would automatically consider withdrawal from the CTBT. In this sense, the President offered deference to the scientific and military leaders. In addition, the ratification package also included the standard "supreme national interest clause," which explicitly reiterated that if the national security interests of the United States were harmed by the treaty, the U.S. could withdraw from the agreement.[23]

The judgments offered at the time by leaders of the U.S. nuclear enterprise clearly reflected the uncertainties about stockpile stewardship but also confidence in its likely effectiveness. In his 1997 letter to Senator Kyl, LANL Director Hecker admitted that since the nuclear testing moratorium "our confidence in the nuclear stockpile has decreased somewhat," adding that "this decline in confidence is an inevitable consequence of lack of testing." At the same time, he also emphasized that "I believe that the SSMP as currently configured and fully funded provides the best approach to keeping the confidence level in our nuclear stockpile as high as possible for the foreseeable future."[24] LLNL Director Tarter voiced a similar conclusion: "Although we have not tested since 1992, I continue to have confidence in the safety and

21 United States Senate, "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

22 Interview with George H. Miller, former Director, Lawrence Livermore National Laboratory (January 22, 2020).

23 Interview with Robert G. Bell, National Security Council Senior Director for Defense Policy and Arms Control (January 1993-August 1999) (April 8, 2020).

24 Hecker, Siegfried S., "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

reliability of the nuclear weapons in the stockpile."[25] His assessment was based on three factors: 1) the weapons in the stockpile were well tested; 2) the United States had a cadre of experienced personnel; and 3) the SSP "puts in place capabilities and methodologies to identify, assess, and respond to problems that occur in the stockpile."[26] Assistant Secretary of Energy for Defense Programs Victor H. Reis echoed the optimism about SSP by stating that "If we stick to what it is supposed to do, which is to maintain the current stockpile indefinitely without testing, I think it can be a substitute. In other words, if you ask it to do that job, it will do that job."[27]

An opposing view was offered by other witnesses. For example, Robert Barker, an assistant to Director Tarter, argued that stockpile stewardship "is not now, and never will be—even 10 years from now when its major components might be operational—a 'substitute' for nuclear testing in the sense of giving us equal confidence in the safety and reliability of our nuclear weapons. Nor will SSP alone allow us to improve the inherent safety of nuclear weapons or provide new nuclear weapon designs in response to new requirements."[28] Others argued that it was much too early to draw definitive conclusions. James Schlesinger, for example, argued that "In assuring weapon reliability, there is no substitute for nuclear testing. How imperfect or how satisfactory a substitute the Stewardship Program will prove to be remains to be seen."[29]

Twenty years later, advocates of testing continue to make their case within the technical community. Writing in 2018, two retired scientists from Los Alamos National Laboratory, John C. Hopkins and David Sharp—both veterans of the pre-moratorium test program—argued that "the scientific foundation for assessments of the nuclear performance of U.S. weapons is eroding as a result of the moratorium on nuclear testing."[30] The core of their argument is that this erosion introduces an unacceptable risk and that a resumption of testing is needed to provide reality checks on SSP and thereby eliminate that risk. Separately, Hopkins published a paper claiming that the "the knowledge needed to conduct a nuclear test, which comes only from

25 Tarter, Bruce C., "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

26 Ibid.

27 Reis, Victor H., "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

28 Barker, Robert B., "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

29 Schlesinger, James, "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

30 Hopkins, John C. and Sharp, David H., "The Scientific Foundation for Assessing the Nuclear Performance of Weapons in the U.S. Stockpile Is Eroding," *Issues in Science and Technology* 35, no. 2 (Winter 2018). https://issues.org/byline/david-h-sharp/. Accessed September 16, 2020.

testing experience, is all but gone" and urged the United States to make nuclear test readiness a priority.[31]

But looking back over 20 years, it's also clear that such views are far from the mainstream in the technical community. A long series of laboratory directors and STRATCOM commanders, supported by large technical teams, have annually certified that the existing arsenal remains safe and effective. This is because they come to different conclusions about the risks involved in the certification program and how to manage them. Thus, it is useful to understand the annual certification process.

In the words of Bruce Tarter, "A critical yearly measure of the success of the SSMP will be our ability to provide formal statements of stockpile confidence through the Annual Certification process."[32] While the certification process is the main factor, Tarter notes that the overall success of the SSP will also depend on indicators like progress within the program, the effectiveness of the tools, and the individual judgements of the scientists and engineers engaged in the SSP.

Despite the testing moratorium and the continued aging of the arsenal, certification has occurred annually. The Department of Energy (DOE)'s National Nuclear Security Administration (NNSA) issued its most recent FY 2020 SSMP report in July 2019, which states that "For the 23rd consecutive year, the science-based Stockpile Stewardship Program has allowed DOE and DoD to certify the safety, security, and effectiveness of the U.S. nuclear weapons stockpile to the President without the use of nuclear explosive testing."[33] This conclusion was echoed in the 2002 and 2012 National Academy of Sciences (NAS) reports as well. The 2012 NAS study found that United States "has the technical capabilities to maintain a safe, secure, and reliable stockpile of nuclear weapons into the foreseeable future without nuclear-explosion testing" and "is now better able to maintain a safe and effective nuclear stockpile and to monitor clandestine nuclear explosion testing than at any time in the past."[34] These SSMP and NAS reports suggest that the majority of the U.S. nuclear technical community finds the stockpile stewardship program a satisfactory alternative to underground testing for the purpose of maintaining the legacy nuclear arsenal and for the associated life extension programs. Indeed, many see it as preferable to nuclear testing, for reasons discussed further below.

To understand the broad confidence of the technical community in SSP, it is necessary to understand precisely how it differs from the old "cut and try"

31 Hopkins, John C., "Nuclear Test Readiness–What is Needed? Why?" *National Security Science*, Los Alamos National Laboratory (December 2016). https://www.lanl.gov/discover/publications/national-security-science/2016-december/_assets/docs/NSS-dec2016_nuclear-test-readiness.pdf. Accessed September 16, 2020.

32 Tarter, Bruce C., "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

33 National Nuclear Security Administration, *Fiscal Year 2020 Stockpile Stewardship and Management Plan*, Report to Congress (July 2019). https://www.energy.gov/sites/prod/files/2019/08/f65/FY2020__SSMP.pdf. Accessed September 16, 2020.

34 National Research Council, *The Comprehensive Nuclear Test Ban Treaty: Technical Issues for the United States* (Washington, DC: The National Academies Press, 2012). https://doi.org/10.17226/12849. Accessed September 16, 2020.

approach to the design, development, and certification of nuclear weapons. The central feature of stockpile stewardship is Uncertainty-Quantified, High-Performance Computing (UQ/HPC). This is a methodology that enables certification without full-scale testing. It is a way of gaining confidence in the performance of a complex system by ensuring it is able to withstand the worst combination of stresses it might ever face. To do so, it adapts the methods of engineering safety factors. Although it may seem that such an approach is unprecedented, in fact many complex systems have been fielded without testing at scale. As further developed in the U.S. nuclear enterprise since its adoption in 2000, the UQ/HPC strategy is essentially engineering safety factors "on steroids." UQ has become the common Livermore and Los Alamos methodology for stockpile certification.[35]

To understand how uncertainty-quantified engineering safety factors can enable certification of a complex design, consider the iconic example of an untestable system—a suspension bridge. No one builds a suspension bridge and tests it to failure. Full-scale testing begins the day the bridge is opened to traffic. Confidence in the bridge's integrity is created in the design phase, which ensures that the resulting structure will be able to safely withstand even the worst combination of insults it may endure in its service life.

Figure 1 illustrates this approach. The point at which the red line crosses out of the yellow box at the left of "normal daily conditions" is the design engineer's best estimate of the worst-case load/stress condition for the bridge. This might include 100 years of corrosion, a loading of the entire span with fully loaded trucks halted in traffic nose to tail in both directions, during a hurricane when a major earthquake occurs. Engineering safety factors should enable the bridge to withstand these combined stresses with sufficient margin-to-failure to spare. How much margin is enough? A competent design engineer is humble, aware that they may not have accounted for, or adequately estimated all of the stresses. Additionally, few suspension bridges have failed so estimates of when failure occurs (i.e. the "knee" in the curve) must have associated uncertainty. Hence, uncertainty (shown as blue blocks) is added to the margin estimates, and there must be plenty of white space (i.e. margin to spare) between the blocks of uncertainty.

Figure 1 applies to any complex system design—including nuclear weapons. After all, these weapons are themselves a complex system of sophisticated technologies designed to perform multiple complex operations in an infinitesimally small period of time: to bring a mass of fissile material to supercriticality, to fission explode, and then to induce a fusion detonation, which then dramatically increases the total yield of the system.

Think of the x-axis in Figure 1 as the nuclear yield of the weapon's first-stage trigger. The y-axis represents the full nuclear yield of the weapon. Clearly, if the trigger

35 Goodwin, Bruce T. and Juzaitis, Raymond J., *National Certification Strategy for the Nuclear Weapons Stockpile* (Livermore, CA: Lawrence Livermore National Laboratory, TR-223486, August 7, 2006). https://pdfs.semanticscholar.org/64c8/39f88acf44346ad2b963fe8327e815e5fb09.pdf. Accessed September 16, 2020.

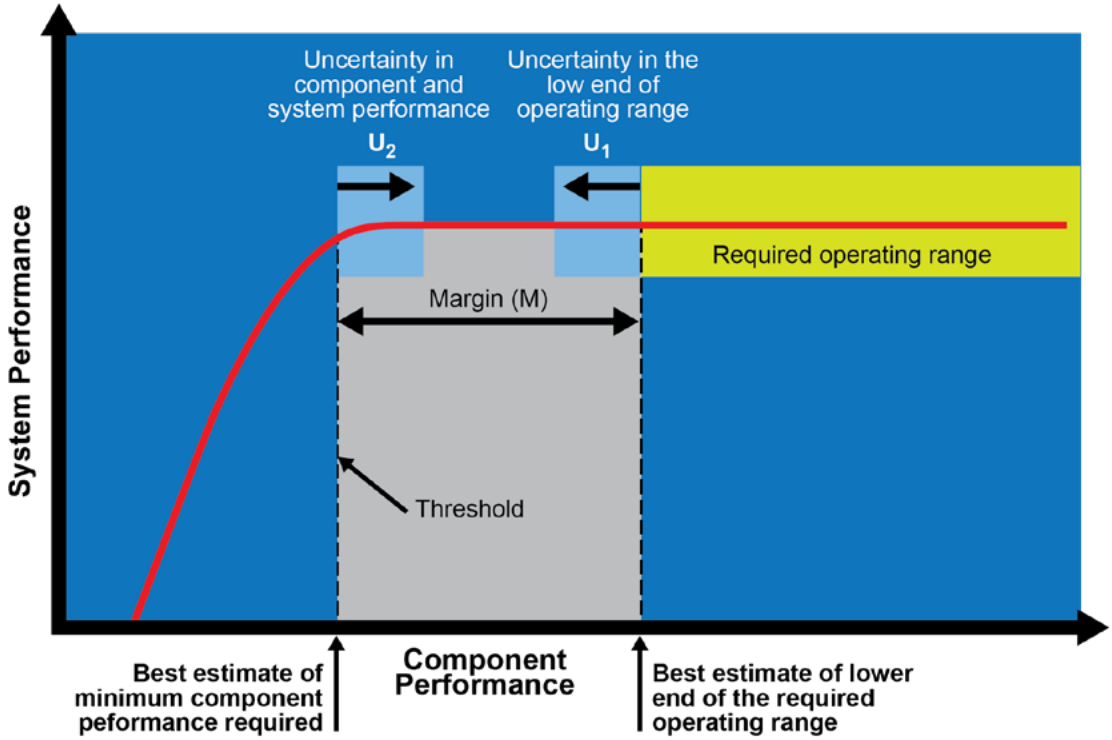


Figure 1. Generic Quantification of Margins and Uncertainty diagram.

yield drops too low, the weapon will fail. The principles of uncertainty quantification analogously apply. The normal operating range of the weapon must have its worst case represented by the left side of the yellow block, with a large margin to failure between that situation and the place where the weapon is likely to begin failing (i.e. the "knee" of the curve). For U.S. nuclear weapons, we have more than 1,000 nuclear test results that define the actual performance curves schematically represented in Figure 1. By assuring significant margin and by accounting for uncertainty, we can certify nuclear weapons without having to resort to nuclear testing.

To illustrate the values of SSP relative to the values of underground testing for sustaining the U.S. nuclear deterrent, let's consider two case studies. The first case involves a particular weapon developed decades ago but kept in service long past its original design life. As was common at the time, while in development it was tested multiple times to ascertain its reliability under different conditions of deployment and age. This weapon catastrophically failed one of these tests. But the test didn't explain why. The cut-and-try method could tell you if a system worked or failed but it would not necessarily tell you what had failed and why. To address the test failure in the pre-SSP era, designers chose an engineering solution that raised the yield to ensure a successful subsequent nuclear test. The change did not, however, fix the basic flaw in the weapon. This resulted in a problematic weapon system that was finally fixed years later when the tools of stewardship enabled the fundamental problem to be accurately determined, simulated, non-nuclear test validated, and subsequently corrected.

In the second case study, a problem that existed for decades in our understanding of weapon performance was resolved via the tools of stewardship. This problem revealed itself in nuclear tests as an apparent violation of a basic law of physics. But, of course, this is not possible. This revealed that the designer's understanding of a very important process in nuclear weapon operation was deeply flawed. The problem could not be solved through nuclear testing, only measured, and so accounted for, under different conditions. The problem was finally resolved by a member of the new generation of stewards of the stockpile. By using the novel tools of stewardship, non-nuclear, hydrodynamic testing at facilities such as the CFF and DARHT, and high-energy density experiments at facilities such as NIF—and combined with ultra-high resolution and massively parallel computer simulations—the underlying physical source of this problem was revealed and so now quantitatively accounted for in weapons operation. As a result, it will never again be necessary to do a nuclear test to measure this specific problem in any nuclear weapon because stewardship has enabled us to properly understand and account for the basic physics underlying the problem.

The lesson here is that the new non-nuclear, experimental measurement capabilities of the SSP reveal and quantify the fundamental processes that must occur for a weapon to work correctly. Using these novel capabilities, the new generation of weapons designers is able to carefully validate the accuracy of simulations of the operation of weapons, and so correctly determine success or failure. They can therefore understand the condition of stockpile weapons across deployment modes as they age, and thereby anticipate failure, and design life extension programs that ensure that U.S. nuclear weapons remain safe, secure, and reliable.

In sum, two decades of experience have removed uncertainty about the efficacy of stockpile stewardship as a substitute for underground nuclear testing as a way to ensure confidence in that the legacy stockpile remains safe and reliable. The risks associated with an aging arsenal and absence of explosive testing are manageable and acceptable because of advanced diagnostic and analytic techniques. An explosive test might add new data but would add little to the confidence gained with those advanced techniques. Indeed, it would be a very poor substitute. These conclusions help explain why the technical community broadly supports SSP and broadly rejects the cases made by Hopkins and Sharp.

These conclusions were already evident in the 2012 NAS study, which judged that the stockpile stewardship program "has been more successful than was anticipated in 1999" and "there is no evidence of any technical issues that cannot be resolved with the present competency."[36] What was cautious optimism by laboratory directors in the 1990s turned into a proven success story.

---

36 National Research Council, *The Comprehensive Nuclear Test Ban Treaty: Technical Issues for the United States* (Washington, DC: The National Academies Press, 2012).

Moreover, stockpile stewardship has provided some real benefits over nuclear testing. As former Energy Secretary Ernest Moniz said in 2015: "The lab directors today now state that they certainly understand much more about how nuclear weapons work than during the period of nuclear testing and, if we had continued the paradigm, what we would understand today in terms of the physics of these devices."[37]

These conclusions have been echoed elsewhere. In 2014, the Congressional Advisory Panel on the Governance of the Nuclear Security Enterprise also emphasized that science-based stockpile stewardship has not only substituted for nuclear testing but actually surpassed it. "To date, Science-Based Stockpile Stewardship has succeeded in sustaining confidence in the U.S. nuclear deterrent. Unmatched technical innovation on the part of NNSA's scientists and engineers has produced a dramatically increased understanding of the country's aging nuclear weapon stockpile."[38]

Or in the words of Parney Albright, former director of Lawrence Livermore National Laboratory: "Testing is unnecessary given the legacy of the components and our understanding of the relevant science (and the significant margins we build into the weapons)."[39] Albright went on to argue that "testing is also a bad strategy, since our adversaries have not made the investments in SSP that we have made, hence have less confidence in their weapons than we do, and hence stand by far the most to gain with a resumption of testing."[40]

An important footnote to this analysis of continued certification of the legacy stockpile relates to material aging. It would not have been possible to maintain the legacy arsenal, to say nothing of certifying its safety and reliability, had the fissile materials aged in a manner that impaired their functionality. The issue arises from the fact that metals exposed to radiation can accumulate microscopic damage when energetic particles emitted from radioactive decay strike atoms in the metal. They can dislodge the atoms from their position in a crystal of metal. These dislodgements are referred to as dislocations. Over time, as the number of dislocations accumulate, the material properties of a metal can change. The concern has been that this process in plutonium could cause changes that might degrade the operation of a nuclear weapon as it ages. Fortunately, it has been found that pits are long lived. How long-lived remains, however, a matter of research and debate.[41]

37 Moniz, Ernest, "Opening Remarks at the NNSA Stockpile Stewardship Program 20th Anniversary Event" (October 22, 1015). https://www.energy.gov/articles/opening-remarks-nnsa-stockpile-stewardship-program-20th-anniversary-event-delivered. Accessed September 16, 2020.

38 Augustine, Norman and Mies, Richard, *Interim Report of the Congressional Advisory Panel on the Governance of the Nuclear Security, U.S. Senate Armed Services Committee* (April 9, 2014). https://www.armed-services.senate.gov/imo/media/doc/Augustine-Mies_04-09-14.pdf. Accessed September 16, 2020.

39 Interview with Parney Albright, former Director, Lawrence Livermore National Laboratory (June 10, 2020).

40 Ibid.

41 Lawrence Livermore National Laboratory, *Manufacturing Plutonium Pits for the Modern Nuclear Stockpile* (Livermore, CA: LLNL, April 21, 2020).

Ensuring the Future Capability to Certify New Weapons

The second Senate technical concern was a close parallel to the first. In 1991, the United States ceased producing nuclear weapons and thus primary attention focused on maintaining the safety, security, and reliability of the legacy arsenal. But it was unclear how long the legacy arsenal might be maintained (the discussion of life extension programs began a decade later). In the 1999 debate, many assumed that the United States would need to return to the design and certification of new nuclear weapons at some unknown future time. Thus, the Senate logically asked whether underground explosive nuclear testing would be necessary for that purpose or whether stockpile stewardship would be an effective substitute. In Senator Kyl's formulation, the central question was: "If U.S. leadership requires a new nuclear design, would you be willing to certify and deploy it without testing?"[42]

In this regard, it is important to note that even before the 1992 moratorium, nuclear testing has not been the only factor in the certification of new nuclear designs. In the words of George Miller, a former LLNL director, "nuclear weapons were never certified by nuclear tests; nuclear tests were important, but frequently not even the most important part of the process because there were never enough nuclear tests over the full range of conditions to provide certification based on the empirical data from those tests. Certification was a statement of confidence and the judgment of technical experts based on a rigorous process that considered all the available data, computational simulations, considerations of margins, etc."[43] Following Miller's logic, the lack of nuclear testing does not automatically mean that a new warhead design cannot be certified.

At the time of the 1999 CTBT debate, Paul Brown, then the assistant associate director for arms control in the National Security Directorate at LLNL, and Thomas Scheber, the project leader for stockpile studies at LANL, launched a collaboration to investigate what the CTBT would mean for the DOE and whether the laboratories' activities were keeping up with the spirit of the treaty. They concluded that the CTBT—as it was explained by the State Department's article-by-article analysis—did not ban the development of new weapons or the improvement of existing warheads. Therefore, the work undertaken by the national labs and any future modernization of the stockpile would not be a violation of the CTBT. The authors also provided a technical definition of the term "new" nuclear design. They argued that any existing

42 Kyl, Jon, "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

43 Miller, George in Brown, Paul, *The Comprehensive Test Ban Treaty: Lawrence Livermore National Laboratory's Impact on U.S. Nuclear Policy from 1958 to 2000* (Livermore, CA: Lawrence Livermore National Laboratory, 2019).

warhead that was refurbished or remanufactured should not be considered a new design. At the same time, a warhead that was not previously in the stockpile and was developed with an untested physics package (i.e. the actual nuclear explosive) should be considered a new design. Brown and Scheber also noted that the problem with these definitions is that in reality there is a lot of room between the above described cases, and the gray zone between these two extremes allows different interpretations of what is actually a new design.[44]

In 1999, the technical advocates of SSP argued that testing would not be necessary—that advanced diagnostic techniques, uncertainty quantification, and the accumulated data from the legacy test program would suffice to enable certification of new designs and their manufacture without underground tests. Others argued that a potential future need for testing could not be ruled out—and might be necessary. In his response to Senator Kyl, Tarter emphasized that the response depends on the design: "I believe it unlikely that an entirely new warhead, developed without the benefit of nuclear testing, would be certifiable by today's standards. However, some modifications of designs that had been previously tested successfully may be possible."[45] In his testimony to Congress, Director of Sandia National Laboratories, Paul Robinson emphasized that "While the treaty does not prohibit the deployment of new nuclear designs, from a practical standpoint we are limited to previously tested concepts." However, he also noted that "many previously tested designs could be weaponized to provide new military capabilities...Proven designs of lower yield exist that might be adaptable for new military requirements in the future. I believe that such weapons could be deployed this way, without the need for nuclear tests."[46]

The Senate's review of this issue was complicated by the fact that the Clinton administration had committed in 1995 to support a treaty that precluded tests of any yield, even very small ones. It posed a challenge for the United States as it contemplated the possibility of future new weapons. In the past, the United States had developed techniques to enable the certification of weapons with tests generating sub-kiloton explosive yields. Now, a new approach was needed. This became another incentive for the development of SSP. The Advanced Simulation and Computing Initiative has been particularly valuable for this purpose.

Today, the technologies and skills developed in the stockpile stewardship program to support the legacy stockpile are closely aligned with the requirements of supporting the development of new weapons. In 2015, Scheber and John R. Harvey summarized the achievements of the past two decades in a National Institute for Public Policy report. The authors emphasized that despite the ongoing testing

44 Brown, Paul, *The Comprehensive Test Ban Treaty: Lawrence Livermore National Laboratory's Impact on U.S. Nuclear Policy from 1958 to 2000* (Livermore, CA: Lawrence Livermore National Laboratory, 2019).

45 Tarter, Bruce C., "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

46 Robinson, Paul, "National Security Implications of the Compreheåsive Test Ban Treaty" (October 6-7, 1999). Hearing before the Committee on Armed Services, S. Hrg. 106-490.

moratorium, every administration since President Clinton required the nuclear enterprise to maintain the necessary personnel and infrastructure that is capable to design, develop, and field new warheads if the need arises. "The intent to sustain a development capability for new warheads is implicit in the maintenance of a Stockpile Stewardship and Management Program."[47] Although the laboratories face many challenges in maintaining this capability, Scheber and Harvey provided a number of recommendations to address these issues, starting with the reform of the training and methods that the next generation of nuclear weapon designers and engineers rely on. In the past three years, NNSA has increased its attention on the issue. As noted in the FY 2020 SSMP, "detailed design, development, qualification, production, and certification of a prototype nuclear explosive package (NEP) are capabilities that are not exercised fully to respond to future warhead requirements. DOE/NNSA established collaborations with the Office of the Secretary of Defense (Nuclear Matters), USSTRATCOM Commander, and relevant Air Force and Navy organizations and began technical work to execute this program in FY 2017. After reviewing emerging threats, technical challenges, and opportunities, DOE/NNSA has selected a set of challenge scenarios for concept studies and potential design, prototyping, and flight testing."[48]

Overall, the United States continues to maintain the capability to design new warheads if there is a need for them, and it has put more emphasis on exercising the necessary skills. Returning to Senator Kyl's original question, the second issue is the certification of these new designs. Twenty years after the CTBT debate in the Senate, the United States is much better positioned to respond to this question. The issue of new-design certification was a prominent question during the Senate debate about the Reliable Replacement Warheads (RRW) in the mid-2000s. The RRW program aimed to create new-design replacement warheads that "will be designed to provide more favorable performance margins than those currently in the stockpile, and will be less sensitive to incremental aging effects or manufacturing variances."[49] The program supporters also emphasized that "RRW designs will be firmly based on previous nuclear test experience."[50] NNSA believed that this new approach would "increase the reliability, safety, and security of the United States nuclear weapons stockpile."[51] The Nuclear Weapons Council, representing the leadership of NNSA and DoD, stated

47 Scheber, Thomas and Harvey, John R., *Assessment of U.S. Readiness to Design, Develop and Produce Nuclear Warheads: Current Status and Some Remedial Steps* (Fairfax, VA: National Institute Press, 2015). https://www.nipp.org/wp-content/uploads/2015/10/Assessment-of-US-Readiness-for-web.pdf. Accessed September 16, 2020.

48 National Nuclear Security Administration, *Fiscal Year 2020 Stockpile Stewardship and Management Plan*, report to Congress (July 2019).

49 U.S. Department of Defense and Department of Energy, *Report on the Feasibility and Implementation of the Reliable Replacement Warhead Program: Submitted to the Congressional Defense Committees Pursuant to Section 3111 of the National Defense Authorization Act for Fiscal Year 2006, P.L. 109-163, by the Secretary of Defense and the Secretary of Energy in Consultation with the Nuclear Weapons Council* (April 2008).

50 Ibid.

51 Ibid.

their confidence “that the RRW strategy is feasible and appropriate for sustaining the Nation’s nuclear warhead stockpile without requiring a return to underground nuclear testing.”[52] Although the program was eventually defunded by Congress, it demonstrated the confidence of the national laboratories that new designs based on the existing test base can be certified without recourse to nuclear testing. Possible future new designs un-related to the test base, were they to be seen as necessary, would be more difficult to certify in this manner.

52 Ibid.

Maintaining Test Expertise Without Testing

The Senate was also concerned with the possibility that the U.S. ability to return to nuclear testing at some future time would atrophy if the CTBT were adopted. As Senator Kyl put it, "How difficult is it, technically, to maintain the capability to test without testing at some level? If CTBT enters into force for the U.S., the budgetary and political pressures to close the NTS [Nevada Test Site] will increase significantly. How important is the retention and maintenance of the NTS?"[53]

Recognizing the salience of this issue, the Clinton administration had already included among the safeguards a requirement to ensure the capacity to return to testing: "The maintenance of the basic capability to resume nuclear test activities prohibited by the CTBT should the United States cease to be bound to adhere to this Treaty."[54] Subsequently, agreement was reached to maintain the capability to return to nuclear testing within two to three years of a decision to do so. This decision was codified in November 1993 by President Clinton's Presidential Decision Directive/NSC-15: "In order to resume underground nuclear tests, a capability to conduct a nuclear test within 6 months up to FY 1996, and to conduct a nuclear test within 2-3 years after that time will be assumed by the Department of Energy."[55]

In the 1997 hearings on stockpile stewardship, the Senate received the following technical advice from leaders of the nuclear enterprise: maintaining NTS is important as the site will remain a critical element of the SSMP even if the CTBT enters into effect. In response to Senator Kyl, Bruce Tarter argued that "It will be difficult to maintain the capability to quickly return to conducting full-scale nuclear tests. However, we plan to use the NTS to provide essential non-nuclear experimental capability to the SSMP."[56] Siegfried Hecker also added that "It is important to exercise the key skills required for nuclear testing. Otherwise, the time to reconstitute will increase substantially. Right now, we find that most of the key skills are being

53 Kyl, Jon. "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

54 Clinton, William J., "Message to the Senate Transmitting the Comprehensive Nuclear Test-Ban Treaty With Documentation," The White House, Office of the Press Secretary (September 22, 1997).

55 Clinton, William J., *Presidential Decision Directive/NSC-15* (Washington, DC: The White House, Clinton Presidential Libraries, November 3, 1993). https://clinton.presidentiallibraries.us/items/show/12743. Accessed September 16, 2020.
Medalia, Jonathan, *Comprehensive Nuclear-Test-Ban Treaty: Updated "Safeguards" and Net Assessments* (Washington, DC: Congressional Research Service, June 3, 2009). https://www.everycrsreport.com/files/20090603_R40612_579784b675ae0980ab4feddda8eee6d4aee83767.pdf. Accessed September 16, 2020.

56 Tarter, Bruce C., "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

exercised with the subcritical tests at NTS."[57] In the words of former LANL Director Michael Anastasio, "the decision to return to testing is a political decision."[58] The President has the ultimate authority to make this decision, and "DOE/NNSA assumes that a test would be conducted only when the President has declared a national emergency or other similar contingency and only after any necessary waiver of applicable statutory and regulatory restrictions."[59] In line with these assumptions, the United States has worked to maintain the capability and technical expertise to return to nuclear testing over the past 20 years, but challenges remain in funding, equipment, infrastructure and knowledge transfer.

Since President Clinton, no administration has updated the guidance on the timing of the test readiness. However, the Bush administration's 2001 Nuclear Posture Review (NPR) instructed NNSA to shorten the timeframe that is needed to resume testing. According to the 2012 NAS study on the CTBT, NNSA internally aimed to reduce the 2-3 year timeframe to 18 months but Congress declined funding for it and instructed to maintain a 24-month readiness capability.[60] Regarding these readiness timeframes, it is also important to note that "NNSA assessments assume the conduct of only a single test or very short series; NNSA is not maintaining any capability to resume routine, sustained nuclear testing."[61]

In practice, the time needed to execute a nuclear test, should one ever be required and funded, would vary depending on the complexity of the test. NNSA's assessment in 2009 was that it had the capability to conduct a test within the required 18 months, but only if the test needed to meet very limited technical objectives, and all the domestic regulations, agreements, and laws were waived.[62] Mark Martinez, current President of the Management and Operating (M&O) contractor that runs the Nevada National Security Site argues that under the above circumstances, a simple test might be conducted within as little as six months today.[63] At the same time, more complex, sophisticated tests would take a year or more to execute. The 2012 NAS study argued that the "development of a weapon with new military characteristics would take

57 Hecker, Siegfried S., "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

58 Anastasio, Michael, "Directors Q&A," *Los Alamos National Laboratory: National Security Science* (February 2014). https://www.lanl.gov/discover/publications/national-security-science/2014-february/_assets/docs/NSS_FEB2013_print.pdf. Accessed September 16, 2020.

59 National Nuclear Security Administration, *Fiscal Year 2020 Stockpile Stewardship and Management Plan*, Report to Congress (July 2019).

60 National Research Council, *The Comprehensive Nuclear Test Ban Treaty: Technical Issues for the United States* (Washington, DC: The National Academies Press, 2012). https://doi.org/10.17226/12849. Accessed September 16, 2020.

61 Ibid.

62 Ibid.

63 Interview with Mark Martinez, President of the Management and Operating (M&O) contractor that runs the Nevada National Security Site (March 31, 2020).

significantly longer than 24-36 months."[64] Cost would also scale accordingly (note that in the past testing costs were up to $100 million per test in 1992 dollars).

According to the 2012 NAS study, a considerable challenge for test readiness is funding. In the past, both Congress and NNSA failed to pay enough attention to this issue and decided not to request or appropriate the necessary funds, which led to "age-related degradation of physical assets and diagnostic equipment, lack of maintenance, outdated technology."[65]

Besides the issue of funding, test readiness has additional elements, including the test site, the equipment and infrastructure, and the body of specialized knowledge. Most of the needed capabilities are preserved at the Nevada Test Site—or, as it is now called, the Nevada National Security Site. There are pre-drilled bore holes, many recently surveyed, with stemming plans (i.e. hole "plugging" plans to prevent radioactive release) ready for test use, were they to be needed, already prepared. Stemming plans need to be test specific. Having example stemming plans in place will help future containment engineers get it right. The test readiness program maintains documentation on these resources. An example is the Nevada Test Site Inventory Emplacement Hole Source Book, a document written for internal use. The means of lifting and lowering very large test canisters continues to be exercised for non-nuclear experiments such as recent large, non-nuclear chemical explosion verification/detection tests. Tunnels are excavated for other non-nuclear facilities and experiments. The technology for the manufacture of large testing canisters that hold the nuclear explosive and the diagnostic measurement equipment for a test is straightforward and relatively unchanged.[66]

If the United States were to use this infrastructure to resume nuclear testing, it would need to adapt to the vast changes in the associated diagnostic technologies that have occurred over the decades and that would significantly change the conduct of such tests. Any attempt to preserve the outdated technologies of the 1990s would be technically and financially unwise—and also imprudent from the points of view of personnel training and readiness. A future test would not use the old 8" floppy computer disks, optical cameras, or analog oscilloscopes to record the ultra-high-speed signals, nor coaxial cables to transmit those high-speed signals up to the surface in any possible future nuclear test. Instead, it would use modern, high-speed digital systems and high-speed fiber optic cables.

But test readiness is not just a matter of technical infrastructure. The nuclear weapons laboratories have retained a core group of designers knowledgeable in past practices, along with the requisite institutional knowledge. They are also developing a new generation of technical staff at the test site, at the nuclear labs, and at the

---

64 National Research Council, *The Comprehensive Nuclear Test Ban Treaty: Technical Issues for the United States* (Washington, DC: The National Academies Press, 2012). https://doi.org/10.17226/12849. Accessed September 16, 2020.

65 Ibid.

66 Interview with Mark Martinez, President of the Management and Operating (M&O) contractor that runs the Nevada National Security Site (March 31, 2020).

production facilities who are skillful in the development and application of the new digital measurement techniques. However, this generation shift and knowledge transfer is not happening without challenges. The 2012 NAS study found that NTS faced some of the same problems that NNSA experienced, including the "lack of experienced and trained personnel for critical positions."[67] As a result of these challenges, the test site has a much smaller cadre today than in the 1990s.

In sum, there are continued issues and difficulties but despite its adherence to the test moratorium, the United States has managed to maintain the capability and expertise to resume nuclear testing if that was ever deemed necessary. This has been possible with preservation of the test site in Nevada but has required much more: ongoing experimental activities at the non-nuclear testing facilities of the NTS, the development of advanced measurement capabilities by the JASPER facility, related activities at the NTS Subcritical Experiments Facility, and related work at DARHT, CFF, and NIF, among others.

Periodically people ask whether, after two decades without testing, it might be prudent to retire the Nevada Test Site. In 1997, Senator Kyl himself alluded to the possibility that the testing moratorium would eventually increase pressures to close the NTS. However, the answer today is the same as it was then: it would not be prudent to do so. Closing the NTS would bring technical risk—that the United States would not have the means to validate an unexpected fix to an unanticipated technical problem. It would also deprive the President of a powerful means to signal U.S. national resolve. Moreover, the NTS does not simply stand by in readiness for a future call to action. It is heavily engaged in ongoing work with relatively high hazard, high consequence non-nuclear test activities. Examples of such activities are very large chemical explosions or subcritical experiments with plutonium and other radioactive or hazardous materials.[68]

67 National Research Council, *The Comprehensive Nuclear Test Ban Treaty: Technical Issues for the United States* (Washington, DC: The National Academies Press, 2012). https://doi.org/10.17226/12849. Accessed September 16, 2020.

68 Lawrence Livermore National Laboratory, "Big Explosives Experimental Facility." https://wci.llnl.gov/facilities/beef. Accessed September 16, 2020.
Department of Energy, "August 30, 2006: Subcritical Test at NTS." https://www.energy.gov/management/august-30-2006-subcritical-test-nts. Accessed September 16, 2020.

Verifying Compliance

An enduring measure of merit for any treaty undergoing U.S. Senate ratification review is whether the United States can reasonably expect to be able to detect militarily significant non-compliance in a way that enables it to respond in a timely manner to safeguard its interests. The verification of compliance is a political process drawing on technical information provided by data declarations and monitoring mechanisms, including both national technical means (NTM) and other means agreed upon by treaty signatories. The expected reliability of the CTBT in ensuring the needed compliance, and of its monitoring provisions in providing the needed timely technical information, were hotly contested in the 1999 debate. A critical factor was the decision of the Clinton administration to agree to the "zero-yield" version of the ban. During the Senate hearings, Senator Kyl wanted to hear from the scientists "What is the U.S. capability, by whatever means, to detect very low-level tests or experiments?"[69]

Some context is needed to understand this debate. Historically, the U.S. government had supported a step-by-step approach of gradually reducing the yield of nuclear tests, putting numerical limits on higher yield testing, getting a handle on underground testing, building a tight verification regime, and continuing to learn how to use science and technology instead of testing to be able to pursue a reduced threshold.[70] This step-by-step process was a continuation of former test ban treaty arrangements and built on the achievements of the LTBT, TTBT, and PNET. The end of the Cold War and the dissolution of the Soviet Union put a ticking clock on this gradual approach, and international pressure was gaining momentum to conclude a multilateral CTB. The George H. W. Bush administration focused on the possibility of a very low threshold test ban treaty, as had the Carter administration in the trilateral CTB talks with the USSR and the U.K. During the treaty negotiations, one of the design parameters of the International Monitoring System (IMS) was to be capable of detecting explosions down to 1 kiloton of yield.

When the Clinton administration came to the White House in 1993, National Security Advisor Anthony Lake launched an inter-agency process to review the U.S. position on nuclear testing and the CTB.[71] The administration concluded that there would be benefit in turning the existing test moratorium into a legally binding and verifiable global ban, not least because it would help to secure the indefinite extension

69 Kyl, Jon, "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

70 Interview with Ron Lehman, director of the Arms Control and Disarmament Agency (1989-1993) (January 16, 2020).

71 Interview with Robert G. Bell, National Security Council Senior Director for Defense Policy and Arms Control (January 1993-August 1999) (April 8, 2020).

of the Nuclear Non-Proliferation Treaty (NPT) in 1995 (then an open question). The inter-agency process first needed to settle on a negotiating venue. Two options were examined: an ad hoc U.S.-led multilateral format, or the Conference on Disarmament (CD). Federal departments were split on this. Despite initial worries about the CD, the success of the Chemical Weapons Convention (CWC) negotiations convinced the different stakeholders to go to the CD. The inter-agency process also decided that maintaining P5 solidarity on the future treaty was essential.

By April 30, 1993, the inter-agency process came to a unanimous recommendation to pursue a 1 kiloton limit and an initial 10-year duration for the CTBT, with a provision for extension. All stakeholders, including the Department of Defense (DoD), the Department of State, the intelligence community, the Department of Energy (DOE), STRATCOM, and the Arms Control and Disarmament Agency (ACDA)[72] were confident that a 1 kiloton threshold would be verifiable. Besides, putting a 10-year limit on the CTBT was meant to give scientists enough time to determine whether the stockpile stewardship program would work. The inter-agency group wanted a verifiable CTBT that would therefore have a good chance to be ratified in the Senate. They believed that finite thresholds served this purpose.

On the morning of the Principals Committee (PC) meeting on the CTBT, the *Washington Post* broke the story of the 1 kiloton limit.[73] This triggered harsh reactions from the Senate (especially from those senators who sponsored the moratorium resolution) as well as anti-nuclear non-governmental organizations, who painted it as a sell-out of the disarmament project and an unnecessary and unwise continuation of the arms race. The public furor created a completely new environment for the PC review. From the DOE, Secretary of Energy Hazel O'Leary participated at the meeting. During the discussion, she pushed back against the 1 kiloton recommendation and advocated for a zero-yield CTBT. As DOE was responsible to maintain the stockpile, Chairman of the Joint Chiefs Colin Powell and the other DoD representatives chose to defer to the DOE position. For the inter-agency process this meant a return to the table to put together a safeguards package that would somehow address the verification challenge of a zero-yield CTBT. This result came as a surprise to many in the administration. After all, the inter-agency process had lined up on the 1 kiloton limit until the last-minute intervention of Secretary O'Leary. The inter-agency process spent the next two years trying to get the zero-yield package together in a form and with assurances sufficient to ensure that it was negotiable and ratifiable.[74]

Meanwhile, Assistant Secretary of Defense Ashton Carter and General Wesley Clark, Director of Strategic Plans and Policy at the Joint Chiefs of Staff (JCS), continued to

72 ACDA was an independent government agency between 1961–1999. The main mission of ACDA was to support the arms control, non-proliferation, and disarmament efforts of the government by developing and implementing policies, strategies, and agreements.

73 Smith, Jeffrey R., "White House Studies Nuclear Test Limits," *Washington Post* (April 30, 1993).

74 Interview with Robert G. Bell, National Security Council Senior Director for Defense Policy and Arms Control (January 1993-August 1999) (April 8, 2020).

advocate for a 1 kiloton limit. O'Leary tasked Assistant Secretary of Energy for Defense Programs Victor H. Reis to get DoD on board with the zero-yield threshold.[75]

With an eye to trying to resolve the technical aspects of this debate, Reis recommended to STRATCOM Commander Admiral Henry G. Chiles to host a workshop in June 1995 that came to be called the "Confidence Conference."[76] The result was again a surprise.[77] Although DoD representatives reiterated their commitment to the 1 kiloton limit, O'Leary had the support of Sandia National Laboratory (SNL) Director Al Narath, who stated earlier that SNL did not need testing to maintain the U.S. stockpile.[78] As Tarter describes in his book *The American Lab: An Insider's History of the Lawrence Livermore National Laboratory*, the laboratory directors "were in a complicated situation because we wanted very much to support Reis and the stewardship program, but as technical people we could not guarantee that it would be successful."[79] All of them were making the calculations between the losses and gains at various yield levels. Paul Brown's book on nuclear testing reveals that Los Alamos Director Siegfried Hecker went into the meeting "wanting some level of testing, hopefully 1 kt. [...] LANL's minimum position was for hydronuclears, anything with a few tons of yield."[80] In the end, the Confidence Conference did not resolve the tensions between the DoD and DOE positions, and the discussions continued through follow-up meetings in Washington, DC. In early August, the laboratory directors got a phone call from DOE headquarters, and they all confirmed that they could endorse a zero-yield decision as long as safeguards and a very robust science-based SSP were put into place.[81] Two days after these phone calls, President Clinton announced publicly that his administration supports a zero-yield CTBT.[82]

Shifting from a 1 kiloton threshold to a zero-yield ban presented many complications. Domestically, this immediately generated criticism from treaty skeptics, heightening concerns about future Senate support. It also sidelined the interagency process at a critical moment—when the requirement emerged for a precise technical definition of the term "zero-yield." None existed at the time, in part because the answer is not as obvious as it might seem. In the end, that definition

75 Interview with Victor H. Reis, former Assistant Secretary of Energy for Defense Programs (February 13, 2020).

76 Despite the initial plan to have a detailed discussion on the meaning of zero-yield, the STRATCOM Confidence Conference had a very broad agenda that included all CTBT-related matters (e.g. hydro-nuclear tests), as well as issues related to stockpile stewardship.

77 Interview with George H. Miller, former Director, Lawrence Livermore National Laboratory (January 22, 2020).

78 Interview with Victor H. Reis, former Assistant Secretary of Energy for Defense Programs (February 13, 2020).

79 Tarter, Bruce C., *The American Lab: An Insider's History of the Lawrence Livermore National Laboratory* (Baltimore, MD: Johns Hopkins University Press, 2018), p320.

80 Brown, Paul, *The Comprehensive Test Ban Treaty: Lawrence Livermore National Laboratory's Impact on U.S. Nuclear Policy from 1958 to 2000* (Livermore, CA: Lawrence Livermore National Laboratory, 2019), p113-114.

81 Ibid. p116.

82 Clinton, William J., "Remarks Announcing a Comprehensive Nuclear Weapons Test Ban," The White House (August 11, 1995). https://www.govinfo.gov/content/pkg/WCPD-1995-08-14/pdf/WCPD-1995-08-14-Pg1432-2.pdf. Accessed September 16, 2020.

was generated by diplomats at the Department of State rather than scientists at the Department of Energy. The Department of State defined the zero-yield obligation as follows: "This means that the agreement prohibits all nuclear explosions that produce a self-sustaining, supercritical chain reaction of any kind whether for weapons or peaceful purposes."[83]

In the subsequent ratification debate, the zero-yield decision featured in two ways. It sharpened concerns about the verifiability of the treaty. It also introduced a new set of questions about the implications of potentially asymmetric obligations being assumed by the United States in a situation in which it was not clear that the other nuclear-weapon-states were also committed to the zero-yield interpretation of the ban. The first of these topics is the focus of the remainder of this section; the second is the focus of the following section.

Regarding the verifiability of a zero-yield ban, the Senate debate was joined on three main points. Could IMS ensure detection of noncompliant activities at test sites generating yields at and below 1 kiloton? Second, could noncompliant activities of a militarily significant kind be hidden by so-called de-coupling means? Third, would hydro-nuclear tests provide militarily significant advantages?

It is important to note that the Senate conducted some of its review behind closed doors to safeguard information deemed secret. These three questions all touch on such matters. The performance of U.S. national technical means as a complement to IMS is not publicly known. Nor are intelligence community assessments about the risks of de-coupling and of the specific military advantages that specific countries might gain with hydro-nuclear testing. This fact of life constrains the ability of analysts working only with unclassified information to argue that their conclusions are definitive.

Could IMS ensure detection of noncompliant activities at test sites generating yields at and below 1 kiloton? In 1999, IMS was only partially complete but it promised good coverage of some regions.[84] Proponents of the CTBT emphasized that the U.S. ability to monitor nuclear tests globally would be increased under the CTBT: "Our ability to monitor the treaty will be enhanced by access to the more than 300 monitoring stations that will make up the CTBT's international monitoring system, and the CTBT requirement for installation of 17 monitoring stations in the Middle East, Lebanon, and China, 31 in Russia will improve our ability to verify this treaty."[85] Although there is no officially stated detection threshold for IMS, it was designed to detect with high confidence an explosion generating 1 kiloton of yield. In the words of Bruce Tarter, "If the proposed seismic, hydroacoustic, low-frequency sound, and radionuclide network of the International Monitoring System (IMS) are installed and

83 Bureau of Arms Control, Verification, and Compliance. "Scope of the Comprehensive Nuclear Test-Ban Treaty," U.S. Department of State (2013). https://2009-2017.state.gov/t/avc/rls/212166.htm. Accessed September 16, 2020.

84 Interview with Jay Zucca (May 7, 2020).

85 United States Senate, "Final Review of the Comprehensive Nuclear Test Ban Treaty" (October 7, 1999). Hearing before the Committee on Foreign Relations, S. Hrg. 106-262.

operated as planned, the system is expected to detect, locate, and identify with high confidence[86] non-evasive[87] explosions with yields of about one-kiloton or above conducted underground, underwater, or in the atmosphere."[88] At the 1999 Senate hearing, Richard Garwin also added that "The IMS would have a good probability of detecting a nuclear explosion anywhere in the world at a level of 1 kiloton. And in many portions of the world, the detectability is much better."[89] The normal practice for calculating the detection threshold magnitude for a network is that detection at three stations with good coverage of an event gives a 95% confidence level for event reporting.[90]

On verification, the White House emphasized two main points during the ratification debate. First, it argued that compliance depends on more than just the ability of U.S. national technical means and the CTBT's IMS to generate timely data. It also depends on the level of uncertainty in the minds of potential violators that they could avoid getting caught. In other words, compliance depends on both detection and deterrence. In the words of Secretary of State Madeleine Albright, "With no treaty, other countries can test without cheating and without limit." She went on to argue that the CTBT would improve U.S. ability to deter and detect clandestine nuclear weapons activity in three ways: intrusive monitoring, comprehensive international verification, and peer pressure: "The more countries that support and participate in the treaty, the harder it will be for others to cheat and the higher the price they will pay if they do."[91]

The second argument by the Clinton administration was that some forms of cheating were more militarily significant than others—and that the cheating most likely to be significant for the United States was also the most detectable. In the words of Senator Kerrey, Vice Chairman of the Senate Committee on Intelligence, "The United States has today the capability to detect any test that could threaten our nuclear deterrence. The type of test that could be conducted without our knowledge could only be marginally useful and would not cause a shift in the existing strategic nuclear balance."[92] The United States, argued the administration, would be more threatened by the covert development of new nuclear weapons by entry-level regional adversaries than by the covert development of improved supplemental capabilities by advanced

86 Bruce Tarter's explanation: "High confidence" is not precisely defined, but here I have in mind the often-used measure of 90%.

87 Bruce Tarter's explanation: An "evasive" test is one that is designed to produce smaller or altered signals, or take advantage of masking by non-nuclear events.

88 Tarter, Bruce C., "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

89 Garwin, Richard, "Final Review of the Comprehensive Nuclear Test Ban Treaty" (October 7, 1999). Hearing before the Committee on Foreign Relations, S. Hrg. 106-262.

90 Interview with Jay Zucca (May 7, 2020).

91 Albright, Madeleine, "Final Review of the Comprehensive Nuclear Test Ban Treaty" (October 7, 1999). Hearing before the Committee on Foreign Relations, S. Hrg. 106-262.

92 Kerrey, Robert, "Final Review of the Comprehensive Nuclear Test Ban Treaty" (October 7, 1999). Hearing before the Committee on Foreign Relations, S. Hrg. 106-262.

nuclear powers. And entry-level programs would be less likely to be capable of successfully masking their activities.

Today, the monitoring system is more robust than in 1999. IMS is currently around 90% complete.[93] In certain regions, it is able to detect explosions with lower yields, as its actual effectiveness is a function of the specifics of geography and geology. The 2012 NAS study found that with the IMS system "the detection probability in regions of better propagation (parts of Asia, Europe and North Africa) is expressed as 90 percent confidence at 90 tons (0.09 kilotons)."[94] However, there are stations that are still missing, particularly in the Middle East. The stations in Iran are installed but turned off. The stations in China have only come online in the last few years.[95] In the absence of entry-into-force of the treaty, the arguments of the Clinton administration that the combination of detection and deterrence provides adequate guarantees for enforcement have not been put to the test.

The second verification issue in the Senate debate was whether noncompliant activities of a militarily significant kind can be hidden by so-called de-coupling means. In the 1999 Senate debate, Senator Kyl argued that "Decoupling is a well-known technique and is technologically simple to achieve. In fact, it is quite possible that Russia and China have continued to conduct nuclear testing during the past seven years, while the United States has refrained from doing so. They could have done so by decoupling."[96]

IMS detects underground explosions by monitoring seismic shock waves moving through geologic structures. Those shock waves result from the "coupling" of an explosion to those structures as the energy released by an explosion acts upon them. De-coupling involves isolating that energy from those structures. This is done most easily with a very large cavity. Salt domes are ideal locations for decoupling cavities and are relatively common across the globe. Solution mining dramatically simplifies and lowers the profile of mining of such cavities.

The risk that such techniques might be useful to illicit nuclear weapons programs was first set out in unclassified form in 1960 in a paper published by the RAND Corporation.[97] The authors claimed that by placing an explosion in a large underground cavity its apparent yield could be reduced by a factor of 300. Their assessment was

93 Comprehensive Nuclear Test-Ban Treaty Organization, A Goal Within Reach: Annual Report 2018 (Vienna: CTBTO Preparatory Commission, July 2019). https://www.ctbto.org/fileadmin/user_upload/pdf/Annual_Report_2018/English/00-CTBTO_AR_2018_EN.pdf. Accessed September 16, 2020.
Hafemeister, David, "The Comprehensive Test Ban Treaty: Effectively Verifiable," *Arms Control Today* 38, no. 8 (2008), p6-12.

94 National Research Council, "The Comprehensive Nuclear Test Ban Treaty: Technical Issues for the United States" (Washington, DC: The National Academies Press, 2012), p105.

95 Comprehensive Nuclear Test-Ban Treaty Organization, *A Goal Within Reach: Annual Report 2018*.

96 Kyl, Jon, "Comprehensive Nuclear Test-Ban Treaty," Senate Executive Session (October 8, 1999).

97 Latter, Albert L., E. A. Martinelli, J. Mathews and W. G. McMillan, "The effect of plasticity on decoupling of underground explosions," RAND Corporation, RM-2665-AEC (November 22, 1960). https://www.rand.org/pubs/research_memoranda/RM2665.html. Accessed September 16, 2020.

highly controversial. Drawing on a pair of nuclear tests code named Salmon and Sterling, others have described the notional decoupling factor as falling within a range of between 50 and 150, with many experts using 70 as a plausible general guide.[98] These tests were executed in a solution mined cavity in the Tatum Salt Dome in Mississippi.

Divide a notional 1 kiloton explosion by a decoupling factor of 70 and you get an apparent yield of approximately 15 tons. That could not be detected by IMS as originally conceived. But this is easier said than done. While one can in principle decouple by these means, the cavity required would be enormous, i.e. approximately 50 meters in radius. This begs a number of questions. For example, could the construction of such a cavity be hidden? Could the cavity's roof integrity be maintained against collapse? Admittedly, partial decoupling could be achieved with smaller cavities. However, mining even a 25-meter cavity, which might achieve a factor of approximately 10, reduction of a 1 kiloton explosion to an undetectable 100 tons of apparent yield would still be an enormous undertaking.

If a country is able to certify weapons operation at sub-kiloton yield, then full seismic decoupling of such tests could be possible. This is especially troubling given statements in the 1990s by Russian nuclear lab experts that anything "undetectable" is allowed.[99] This would make detection of such a test by IMS problematic. In this regard, the two most important findings of the 2012 NAS study was that "For IMS and open monitoring networks, methods of evasion based on decoupling and mine masking are credible only for device yields below a few kilotons worldwide and at most a few hundred tons at well-monitored locations. The states most capable of carrying out evasive nuclear-explosion testing successfully are Russia and China. Countries with less nuclear-explosion testing experience would face serious costs, practical difficulties in implementation, and uncertainties in how effectively a test could be concealed. In any case, such testing is unlikely to require the United States to return to nuclear-explosion testing."[100]

We note that for tests occurring at test sites, seismic templates to previous tests enable the detection of new explosions today by IMS down to a few tons of TNT yield or less for an underground explosion.[101] Thus, even a decoupled test of a few hundred tons at an identified test site should be detectable by IMS. In light of this, decoupling today is not as simple or assured of evasion as Senator Kyl suggested in 1999.

98 United States Department of Energy, "Salmon, Mississippi, Site," fact sheet (December 2019). https://www.energy.gov/sites/prod/files/2019/12/f69/SalmonFactSheet.pdf. Accessed September 16, 2020.

99 Interview with George H. Miller, former Director, Lawrence Livermore National Laboratory (January 22, 2020).

100 National Research Council, *The Comprehensive Nuclear Test Ban Treaty: Technical Issues for the United States* (Washington, DC: The National Academies Press, 2012), p112.

101 Ford, Sean R. and Walter, William R., "International Monitoring System Correlation Detection at the North Korean Test Site a Punggye-ri with insights from the Source Physics Experiment," *Seismological Research Letters* 86, no. 4 (July/August 2015). https://pubs.geoscienceworld.org/srl/article-lookup/86/4/1160. Accessed September 16, 2020.

We also note that there are many more activities surrounding a nuclear test that may be detected by other national technical means. As the NAS study noted, "NTM [gave] the United States significant additional information beyond what is available to other countries that do not have a robust NTM program. U.S. NTM can focus on monitoring countries of concern to the United States. The U.S. global monitoring capabilities are generally better than those of the CTBTO because they can enhance data available to the CTBTO with classified capabilities."[102] Besides, "humint" (i.e. intelligence reporting from human observations) may also provide key information about states attempting to cheat. However, "humint," unlike seismic detection, is not necessarily within the control of the observing parties and thus less dependable.[103]

Finally, there is the question of hydro-nuclear tests. Can we detect hydro-nuclear tests? These are tests that achieve a sustained supercritical chain reaction and may produce grams to kilograms of fission yield. They therefore violate the criteria for zero-yield. This level of yield is generally much less than the chemical explosive yield that implodes a nuclear weapon. The United States considers such tests to be banned under the CTBT, though others may not. During the Senate hearings, Republican lawmakers argued that Russia sees hydro-nuclear tests as permitted under the CTBT (we will return to this topic in the following section).

Detecting such violations of the treaty is very problematic, if not impossible. IMS clearly cannot detect events more than 100,000 times smaller than it was originally designed to detect. To detect such an event, other outputs of the supercritical chain reaction would have to be observed. There are two other outputs. First, the burst of neutrons emitted by the chain reaction and, second, the fission products of the chain reaction. We presume that the hydro-nuclear experiment would be conducted in an underground cavity to safely prevent the explosive dispersal of fissile materials and fission products. Hence, the detection of the neutron emission is precluded.

IMS has a robust radionuclide monitoring system, but questions remain about its ability to adequately detect low-yield tests. Radionuclide detection makes it possible to detect escaping fission product particulate matter or gases from an underground nuclear explosion. Confirming the presence of radioactive gases or particles confirms that an explosive event was in fact nuclear. One of the most detectable leakage gases is the noble gas xenon (Xe).[104] Noble gases are useful for detecting nuclear tests because they do not react with surrounding soil and earth within the testing cavity. Xenon is a commonly used noble gas for CTBT verification since atmospheric

102 National Research Council, *The Comprehensive Nuclear Test Ban Treaty: Technical Issues for the United States* (Washington, DC: The National Academies Press, 2012), p121.

103 Medalia, Jonathan, *Comprehensive Nuclear-Test-Ban Treaty: Issues and Arguments* (Washington, DC: Congressional Research Service, RL34394, March 12, 2008). https://www.everycrsreport.com/files/20080312_RL34394_5daebde344d5f2bd96981ecce2dda1e08fc2293a.pdf . Accessed September 16, 2020.
Hafemeister, David, *Physics of Societal Issues: Calculations on National Security, Environment, and Energy* (New York: Springer Science & Business Media, 2013), p134.

104 We greatly appreciate the assistance of Megan Cordone, who helped us develop a deeper understanding of the challenges of xenon detection.

background levels of xenon specifically are much lower than other noble gases such as krypton, which still lingers in the atmosphere in high amounts since the age of atmospheric nuclear tests. Xenon radioisotopes that fall within the mass range of 131-135 are useful for detection because they are near the peak of the fission product mass yield curve for uranium and plutonium and are consequently produced in detectable quantities after nuclear detonations.[105] Xenon isotopes in this mass range have half-lives that are typically long enough for IMS stations to detect.

The CTBTO IMS system has very sensitive stations for the detection of fission product leakage and those stations are getting increasingly more sensitive as new detection machines and methods are being developed. The problem with radionuclide detection and specifically radioxenon detection is that particulates are unlikely to get into the atmosphere from underground cavities as even emerging nuclear nations become increasingly better at containing nuclear tests. At the NTS, for example, there was a minimum depth of burial for a failed nuclear test of 600 feet to limit possible radioactive material leakage detectable at the test site boundary. Besides, there is a constantly increasing level of background xenon isotopes from nuclear reactor operations, as well as from medical and industrial sources, especially in the Northern Hemisphere. This masks and can confuse radioxenon detection.[106] When different sources of xenon mix are detected by an IMS station, it is hard to determine where the xenon samples came from and how much of the sample came from a suspected nuclear test versus a medical isotope production source or a nearby nuclear reactor.[107] To make matters worse, ultra-low-yield experiments could, in principle, be confined to a reusable metal vessel—thus preventing any and all emissions of radionuclides in either gas or particulate form. Since the inception of the CTBT, "IMS radionuclide network has gone from being essentially non-existent to a nearly fully functional and robust network with new technology that has surpassed most expectations."[108] This network has proven extremely useful in IMS for decades, but as CTBT evasion techniques become more advanced, xenon leakage is becoming less and less effective as a metric to detect low-yield nuclear tests.

Hence, the best way to verify that there are no supercritical fission transients is to have "someone in the room" during the experiment. Thus "humint" could play a role in verifying compliance with the "zero-yield" criteria.

There is some confusion in the non-technical community on the definitions of hydrodynamic tests (hydro tests for short) versus hydro-nuclear tests. In each type

---

105 Bowyer, Ted W. et al., "Detection and analysis of xenon isotopes for the comprehensive nuclear-test-ban treaty international monitoring system," *Journal of Environmental Radioactivity* 59 (2002), p139-151.

106 Bowyer, Ted W. et al., "Detection and analysis of xenon isotopes for the comprehensive nuclear-test-ban treaty international monitoring system."

107 Bowyer, Ted W. et al., "Maximum reasonable radioxenon releases from medical isotope production facilities and their effect on monitoring nuclear explosions," *Journal of Environmental Radioactivity* 115 (2013), p192-200.

108 National Research Council, *The Comprehensive Nuclear Test Ban Treaty: Technical Issues for the United States* (Washington, DC: The National Academies Press, 2012), p7.

of test, a full weapon device is built up for the experiment. In a hydro test, the weapon fissile material is replaced with a non-fissile surrogate material. Thus, it is not possible to have a supercritical transient or any nuclear yield in a hydro test. In a hydro-nuclear test, the amount of fissile material in the weapon may be reduced so as to prevent significant nuclear yield, but a supercritical transient is possible with associated minor, i.e. non-explosive, nuclear yield.

In the debates surrounding the moratorium on nuclear testing, it is asserted that hydro-nuclear testing is relatively unimportant, especially for advanced nuclear states with significant historical testing experience such as the United States or Russia. Hydro-nuclear experiments are generally used to assess the safety of nuclear weapons in accidents. During the Cold War, both the United States and the Soviet Union conducted hydro-nuclear tests. Thorn and Westervelt in 1987 published a good review of hydro-nuclear testing in the United States.[109] Hydro-nuclear tests are used to experimentally determine if a nuclear weapon will experience a supercritical transient in an accident where the chemical high explosive has been detonated at a single point (i.e. often referred to as the "one-point safety criteria"). Such a detonation is very asymmetric and so unlikely to produce a criticality transient, but safety demands certainty under accident conditions. An accident is quite three dimensional (3D) as the weapon interacts with its surroundings, and this strongly influences its one-point safety behavior. Hence, a simulation of such an accident must be fully 3D and very highly resolved. The capability to do such simulations have been enabled by the advances made in computing by the ASCI of the SSP. Thus, the purposes for hydro-nuclear experiments described in the 1986 Thorn and Westervelt paper have been replaced by the technical advances in computing of the last three decades. The extensive data from past U.S. hydro-nuclear experiments validate the exquisitely resolved 3D simulations that have made hydro-nuclear testing unnecessary for the United States. Since the end of the Cold War, several scientific panels have come to the same conclusion. A 1994 JASON study on stockpile stewardship judged that hydro-nuclear tests were not needed for the United States to maintain its stockpile.[110] In 2012, the NAS study on the CTBT also argued that "Hydronuclear tests would be of limited value in maintaining the United States nuclear weapon stockpile in comparison with the advanced tools of the Stockpile Stewardship Program."[111]

Russia has conducted many more hydro-nuclear experiments than the United States, and these tests were a much more integrated part of nuclear weapons development than in the case of the United States. According to the NAS scientists, in the case of Russia additional experiments could be useful in maintaining their existing

---

109 Thorn, R N, and Westervelt, D R., *Hydronuclear experiments* (LA-10902=MS (U), United States: N, 1987). https://www.osti.gov/biblio/6646692. Accessed September 16, 2020.

110 JASON Group, *Science Based Stockpile Stewardship JSR-94-345* (The MITRE Corporation, 1994). http://www.fas.org/irp/agency/dod/jason/sbss.pdf. Accessed September 16, 2020.

111 National Research Council, *The Comprehensive Nuclear Test Ban Treaty: Technical Issues for the United States* (Washington, DC: The National Academies Press, 2012), p5.

arsenal, "including life extension and modifying materials to improve safety."[112] However, the NAS group judged it unlikely that these tests would help Russia to develop new strategic capabilities. Russia has the ability to develop and deploy new low-yield tactical nuclear weapons based on its past designs, without new nuclear-explosion tests.

China lacks hydro-nuclear test experience from the Cold War period. Therefore, it is unclear how they could benefit from such tests in their current modernization efforts.[113]

In general, since hydro-nuclear tests are of use principally in determining the safety of nuclear explosives, the likelihood of undetectable violations would very much depend upon the concern a proliferating nuclear state had with respect to weapons safety.

In sum, after 20 years, significant questions remain about the efficacy of the CTBT monitoring systems and its reliability in providing timely information about non-compliant behavior by other states. It is likely capable of detecting militarily significant testing activity at known test sites. But various forms of activity, militarily significant and otherwise, may not be detectable, especially if conducted at covert test sites. In addition, states like Russia and China have the capability to conduct tests with evasive techniques that could avoid detection, and a hydro-nuclear test by Russia "fully contained in a properly designed explosive containment vessel would likely reveal nothing to remote monitors."[114]

To put these findings in context, it is useful to recall the argument of the Clinton administration that deterrence and detection are both important—just because an adversary could cheat, this does not mean that it is willing to risk being caught. In this regard, 20 years after the CTBT debate we are still not able to fully explain the extent to which IMS serves as a deterrent to cheating. Also, not knowable without access to classified information is the degree to which U.S. national technical means compensate for deficiencies (from a U.S. perspective) in the IMS monitoring system. We only know that national technical means provide superior capability in comparison to the IMS but this in itself is not enough to come to a definitive conclusion.

112 Ibid. p103-104.

113 Ibid. p104.

114 Ibid. p101.

Competing at Zero Yield

As noted above, the zero-yield decision sparked a debate about whether the United States had put itself at a disadvantage, given doubts about whether the other nuclear weapons states, especially Russia and China, were committed to the same interpretation of the ban, or a similarly strict standard of compliance. Would they be able to gain some advantage, military or otherwise, over time with experiments the United States had foresworn? This issue has recently regained prominence with reports by the Department of State that both Russia and China are engaged in weapons development activities inconsistent with the zero-yield requirement as the United States understands it.[115]

In the negotiations, the United States sought to eliminate any room for different legal interpretations of the treaty. Article I states that “Each State Party undertakes not to carry out any nuclear weapon test explosion or any other nuclear explosion, and to prohibit and prevent any such nuclear explosion at any place under its jurisdiction or control.”[116] The Department of State’s article-by-article analysis of the CTBT clearly states that the United States understands that “the prohibition on carrying out a nuclear weapon test explosion or any other nuclear explosion applies universally,”[117] leaving no loopholes for example to hydro-nuclear experiments with a very low nuclear yield.

But questions arose during the Senate debate about whether Russia and China had the same interpretation of the Article I commitment. Ambassador Stephen Ledogar, chief U.S. negotiator of the CTBT, had already in 1996 set out his understanding of the commitments made by his counterparts from Moscow and Beijing as follows: “I have heard some critics of the Treaty seek to cast doubt on whether Russia…committed itself…to a truly comprehensive prohibition of any nuclear explosion, including an explosion…of even the slightest nuclear yield. In other words, did Russia agree that hydronuclear experiments, which do produce a nuclear yield, although usually very, very slight, would be banned and that hydrodynamic explosions, which have no yield because they do not reach criticality,

---

115 United States Department of State, *Adherence to and Compliance with Arms Control, Nonproliferation, and Disarmament Agreements and Commitments (Compliance Report)* (April 2020). https://www.state.gov/adherence-to-and-compliance-with-arms-control-nonproliferation-and-disarmament-agreements-and-commitments-compliance-report/. Accessed September 16, 2020.

116 United Nations, *Comprehensive Nuclear-Test-Ban Treaty (CTBT).*

117 United States Department of State, *Article-By-Article Analysis of The Comprehensive Nuclear Test-Ban Treaty* (September 24, 1996). https://2009-2017.state.gov/t/avc/trty/16522.htm. Accessed September 16, 2020.

would not be banned? The answer is a categorical 'yes.' The Russians as well as the rest of the P-5 did commit themselves."[118]

Official Russian and Chinese statements of that period aligned with Ledogar's view. Grigory Berdennikov, Russia's ambassador to the Conference on Disarmament, underlined that "Russia accepted that any nuclear weapons-test explosion or any other nuclear explosion in any environment will be banned forever and without any thresholds."[119] Chinese Ambassador Sha Zukang stated that "the CTBT will without any threshold prohibit any nuclear-weapon test explosion."[120]

But doubts remained. During the Senate debates, Senator Jesse Helms argued that the "Russian Government has clearly stated the view that hydronuclear testing is permitted."[121] Treaty opponents argued that: "It is a very simple task for Russia, China, or others to hide their nuclear tests...In fact, it is quite possible that Russia and China have continued to conduct nuclear testing during the past seven years, while the United States has refrained from doing so. They could have done so by decoupling. There are also other means of cheating that can circumvent verification."[122] As discussed in the preceding section, the treaty's verification provisions would be unlikely to detect such testing by Russia or China.

Senator Kyl approached the issue by trying to answer the question of what amount of yield would a clandestine foreign nuclear test be a technically and militarily significant violation of the CTBT. In response, he heard a variety of expert views. In his response, Hecker stated that "it depends greatly on whether the testing entity is an established nuclear power (even here, our response would be different for Russia and China), a rogue state, or a terrorist. We can state that crude nuclear weapons can, with the knowledge that is in the open literature today, be fielded without any nuclear testing. It is instructive to note that the bomb dropped on Hiroshima was not tested."[123] Tarter also emphasized that the response is different for different states: "I am qualified to address the technical significance of violations at various yield levels. The military significance of such violations would best be answered by military experts in the DOD. The technical significance depends strongly on the technological

118 United States Department of State, "Key P-5 Public Statements on CTBT Scope," fact sheet, Bureau of Arms Control and Verification (1996). https://2009-2017.state.gov/t/avc/rls/173945.htm. Accessed September 16, 2020.

119 Ibid.

120 Ibid.

121 United States Senate, "Final Review of the Comprehensive Nuclear Test Ban Treaty" (October 7, 1999). Hearing before the Committee on Foreign Relations, S. Hrg. 106-262.

122 United States Congress, "Comprehensive Nuclear Test-Ban Treaty" (October 8, 1999). Executive Session of U.S. Senate. https://www.congress.gov/congressional-record/1999/10/08/senate-section/article/S12257-6. Accessed September 16, 2020.

123 Hecker, Siegfried S.,"Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

capability of the country performing the test, the type of device, the information they are seeking (e.g., a reliability test), and the uncertainty they are willing to accept."[124]

As Tarter's response suggests, the discussion of militarily significant cheating was a topic for military experts. Given the sensitive nature of those discussions, they were conducted with secret information and thus are not a part of the public record.

In response to these concerns, the Clinton administration made the case that there were no significant advantages to be gained by others who might attempt to violate the CTBT's zero-yield obligation with occasional evasive tests. In the words of Secretary of Defense William Cohen, "We would not be able to detect every evasively conducted nuclear test. But from a national security standpoint, this is not going to be dispositive in my judgement. Indeed, I am confident that the United States will be able to detect a level of testing—the yield and the number of tests—by which a state could undermine the U.S. nuclear deterrent."[125]

In the interim 20 years, this issue has only grown in significance. There have been occasional hints that Russian leadership has not adhered to the zero-yield standard. Russian nuclear design lab officials, for example, have stated that anything undetectable is allowed, which suggests a much broader interpretation of Article I.[126] The issue has come back to the fore with the 2020 State Department Report titled *Adherence to and Compliance with Arms Control, Nonproliferation, and Disarmament Agreements and Commitments* (also known as the Compliance Report). The report stated that the United States has concerns about Chinese adherence to the zero-yield standard, and it has clearly stated that "Russia has conducted nuclear weapons experiments that have created nuclear yield and are not consistent with the U.S. 'zero-yield' standard."[127]

Might Russia and China be gaining somehow with yield-generating testing? Two decades after the CTBT ratification debate, this issue remains unsettled. This has something to do with doubts and uncertainties about the nuclear ambitions of Russia and China. It also has something to do with the fact that much of the relevant information is sensitive national security information and thus not part of the public discourse.

At present, two competing arguments have taken shape. One concludes that Russia and China are gaining something valuable and one concludes that they are not.

The case that they are gaining an advantage runs along the following lines. Both Russia and China are engaged in nuclear modernization campaigns involving warhead improvements. Those improvements are driven by Russia's stated interest in very

---

124 Tarter, Bruce C., "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267.

125 Cohen, William S., "National Security Implications of the Comprehensive Test Ban Treaty" (October 6-7, 1999). Hearing before the Committee on Armed Services, S. Hrg. 106-490.

126 Interview with George H. Miller, former Director, Lawrence Livermore National Laboratory (January 22, 2020).

127 United States Department of State, *Adherence to and Compliance with Arms Control, Nonproliferation, and Disarmament Agreements and Commitments (Compliance Report).*

low-yield options[128] and China's interest in improved yield-to-weight ratios as it moves to deploy multiple warheads atop its missiles.[129] In this context, low-yield tests may allow them to gain confidence in warhead modernization that they would not otherwise have. Moreover, according to the argument, the United States has denied itself such new capabilities with the zero-yield interpretation. For example, the director of the Defense Intelligence Agency, Lt. Gen. Robert P. Ashley Jr., argued in 2019 that "The United States believes that Russia probably is not adhering to the nuclear testing moratorium in a manner consistent with the zero-yield standard. Our understanding of nuclear weapon development leads us to believe Russia's testing activities would help it improve its nuclear weapon capabilities. The United States, by contrast, has forgone such benefits by upholding a zero-yield standard."[130]

In contrast, the case that Russia and China are not gaining advantages rests on the following arguments. First, the new capabilities of Russia and China, while troubling, do not materially affect strategic stability or the balance of nuclear power. Even if new Russian weapons that are very low yield might have some perceived tactical utility for the Russian military, this is not a capability the U.S. military has sought. Even if new Chinese missiles include multiple warheads, they will not deprive the United States of a credible threat of nuclear retaliation. Second, U.S. nuclear modernization can proceed successfully without yield-generating tests, as argued in a prior section of this paper.

In our judgement, the stronger of the two cases made above is the latter. Although Russia and China remain countries of concern with regards to evasive techniques, we share the conclusions of the 2012 NAS study that the military value of these tests would be at most very limited: "Based on Russia's extensive history of hydronuclear testing, such tests could be of some benefit to Russia in maintaining or modernizing its nuclear stockpile. However, it is unlikely that hydronuclear tests would enable Russia to develop new strategic capabilities outside of its nuclear-explosion test experience."[131] Besides, so long as the United States maintains confidence in its existing arsenal and in the certifiability of possible future new weapons, there is no advantage that Russia or China can gain through low-yield testing that would decisively upset the strategic balance.

128 Woolf, Amy, *Russia's Nuclear Weapons: Doctrine, Forces, and Modernization* (Washington, DC: Congressional Research Service, R45861, January 2, 2020). https://www.everycrsreport.com/reports/R45861.html. Accessed September 16, 2020.

129 United States Congress, "Nuclear Weapons Modernization in Russia and China: Understanding Impacts to the United States" (October 14, 2011). Hearing before the U.S. Senate. https://www.govinfo.gov/content/pkg/CHRG-112hhrg71449/html/CHRG-112hhrg71449.htm. Accessed September 16, 2020.

130 Sonne, Paul, "Top U.S. military intelligence official says Russia 'probably' not adhering to nuclear test ban," *Washington Post* (May 29, 2019). https://www.washingtonpost.com/world/national-security/top-us-military-intelligence-official-says-russiaprobably-not-adhering-to-nuclear-test-ban/2019/05/29/815a1a36-8234-11e9-9a67-a687ca99fb3d_story.html. Accessed September 16, 2020.

131 National Research Council, *The Comprehensive Nuclear Test Ban Treaty: Technical Issues for the United States* (Washington, DC: The National Academies Press, 2012), p5.

Conclusions

The technical context in which the strengths and weaknesses of the CTBT was debated two decades ago has changed a great deal in the interim. The national debate should account for this changed context.

Some technical judgments have been validated. These include, for example, the judgments that IMS is capable of detecting most significant violations at established test sites but that the monitoring of very low yield and de-coupled explosions away from established testing sites is likely beyond the reach of IMS.

Many of the major technical uncertainties surrounding stockpile stewardship have been settled. The Stockpile Stewardship Program has proven to be a success. A return to underground testing is not necessary to sustain a safe and effective legacy arsenal, nor even to design new weapons should the need arise. This is illustrated by the stockpile life extension programs that are successfully fixing problems that have arisen in aging weapons systems.

But some important uncertainties remain and are likely to remain for a long time. These include, for example, the long-term consequences of the Russian and Chinese nuclear modernization efforts for the competitive position of the United States.

In addition, some new technical questions have come into focus, such as the trade-offs between the costs of a potential return to underground testing versus the ability over time to both assess the reliability of each weapon and to modernize the stockpile and enterprise. This is a question, in a finite budget, of the opportunity costs of renewed nuclear tests diminishing sustainment and modernization activities.

Appendix 1

Senator Kyl's Questions to the Design Laboratory Directors[132]

In 1997, Senator Kyl sent a list of 21 questions to LANL Director Siegfried S. Hecker and LLNL Director Bruce Tarter. Their written responses became part of the Congressional record during the Senate hearing on the "Safety and Reliability of the U.S. Nuclear Deterrent" (S. Hrg. 105-267). This hearing addressed a wide range of technical issues related to the Stockpile Stewardship Program. Despite the strong connection between the issues of SSP and the ban on nuclear testing, not all of these questions were in the forefront of the 1999 CTBT technical debate. Our analysis only focused on those questions that were prominent during the 1999 hearings. For the sake of completeness, below is the entire list of questions from 1997:

1. Will confidence in the safety and reliability of U.S. nuclear weapons decline without nuclear testing?
2. Do you expect the Stockpile Stewardship and Management Program (SSMP) to give you the same confidence in the stockpile as was achieved by nuclear testing? If not, by how much will confidence be reduced, assuming the SSMP is successful?
3. What proportion of the research and testing envisioned for the first 10 years of operation of the National Ignition Facility (NIF) is directly related to nuclear weapons? What proportion is indirectly related to nuclear weapons?
4. A purpose of SSMP is to maintain a cadre of scientists and technicians who will be capable of designing and working on nuclear weapons. Will scientists and technicians working on SSMP have weapons classification clearances and will they have a clear commitment to working on nuclear weapons should the need arise?
5. Much of the capability of SSMP is a decade or more away from being fully functional. Furthermore, many of the technologies involved are unproven. From a technical standpoint, would it be advisable to conduct nuclear tests to calibrate the existing and planned technologies? If so, what is the lowest yield at which meaningful tests can be conducted? What is the minimum number of tests that would be required in the interim before SSMP becomes fully functional?
6. What are the specific measures by which you will know whether SSMP has succeeded or failed?

132 United States Senate, "Safety and Reliability of the U.S. Nuclear Deterrent" (October 27, 1997). Hearing before the Subcommittee on International Security, Proliferation, and Federal Services, S. Hrg. 105-267. https://www.govinfo.gov/content/pkg/CHRG-105shrg44720/html/CHRG-105shrg44720.htm. Accessed September 16, 2020.

7. Since the last U.S. nuclear test, have there been age-related or other changes in the stockpile that previously would have been addressed by conducting at least one nuclear test? If so, how certain are you of the fixes? If your level of confidence in the fixes is not extremely high, how has this affected your view of stockpile reliability?
8. How safe is the stockpile today? Have there been any changes since the 1990 Drell safety study that would have changed the conclusions of that study today?
9. What known safety vulnerabilities are we accepting? Should we be accepting them?
10. Are there any tests you would advocate doing today, if allowed, to address safety or reliability concerns?
11. If U.S. leadership requires a new nuclear design, would you be willing to certify and deploy it without testing?
12. What yield of testing would be the lowest possible to accomplish new designs as well as safety and reliability.
13. How difficult is it, technically, to maintain the capability to test without testing at some level?
14. If CTBT enters into force for the U.S., the budgetary and political pressures to close the NTS will increase significantly. How important is the retention and maintenance of the NTS?
15. Why did your laboratory change its long-held view that nuclear testing is essential?
16. What is your understanding of the limitations imposed by "zero"? Are these limitations acceptable in your view?
17. What are your major concerns about your ability to fulfill your responsibilities under a zero CTBT?
18. What importance do you attach to being able to exercise the "supreme national interest" test?
19. What is the monitoring capability of the international system? Of U.S. national technical means?
20. What is the U.S. capability, by whatever means, to detect very low level tests or experiments?
21. At what yield would a clandestine foreign nuclear test be a technically and militarily significant violation of the CTBT?

Appendix 2:

Chronology[133]

1946: Baruch Plan on disarmament

In June 1946, the United States introduced the Baruch Plan to the United Nations Atomic Energy Commission. The United States, the United Kingdom, and Canada jointly called for the establishment of an international organization that would regulate atomic energy. The plan called for the elimination of all nuclear weapons, and a ban on nuclear testing.

1958: The first round of CTB negotiations

In August 1958, President Eisenhower made a proposal "to negotiate an agreement with other nations which have tested nuclear weapons for the suspension of nuclear weapons tests and the establishment of an international control system."[134] In October, the United States, the United Kingdom, and the Soviet Union began the first round of trilateral talks in Geneva on a comprehensive nuclear test ban (CTB) agreement. All three states also announced a one-year testing moratorium.

1958-1961: "Gentleman's" nuclear testing moratorium

In the 1958-1961 timeframe, the United States, the United Kingdom, and the Soviet Union extended their voluntary testing moratoriums. Due to growing international tensions, in September 1961 the Soviet Union broke out of the testing moratorium with 135 nuclear tests in 15 months.

1963: LTBT

President Kennedy signed the Limited Test Ban Treaty with the United Kingdom and the Soviet Union on August 5, 1963. The LTBT prohibited nuclear weapons tests "or any other nuclear explosion" in the atmosphere, in outer space, and under water.

1968: NPT

On July 1, 1968, the Treaty on the Non-Proliferation of Nuclear Weapons (NPT) is signed by the United States, the Soviet Union, the United Kingdom, and 58 other

---

133 Kimball, Daryl, "Nuclear Testing and Comprehensive Test Ban Treaty (CTBT) Timeline," Arms Control Association (June 2019). https://www.armscontrol.org/factsheets/Nuclear-Testing-and-Comprehensive-Test-Ban-Treaty-CTBT-Timeline. Accessed September 16, 2020.
Sherman, Robert, "Comprehensive Test Ban Treaty Chronology," Federation of American Scientists (November 2001). https://fas.org/nuke/control/ctbt/chron.htm. Accessed September 16, 2020.

134 Eisenhower, Dwight D., "Statement by the President Following the Geneva Meeting of Experts Proposing Negotiations on Nuclear Controls" (August 22, 1958). https://www.presidency.ucsb.edu/documents/statement-the-president-following-the-geneva-meeting-experts-proposing-negotiations. Accessed September 16, 2020.

countries. In the treaty preamble, there is an explicit reference to a comprehensive test ban agreement, as well as the expressed will of the state parties to "discontinue" all nuclear explosive tests.

1974: TTBT

The Treaty on the Limitation of Underground Nuclear Weapon Tests, also known as the Threshold Test Ban Treaty (TTBT), was signed in July 1974 by the United States and the Soviet Union. It established a nuclear "threshold" by prohibiting tests having a yield exceeding 150 kilotons.

1976: PNET

The Treaty on Underground Nuclear Explosions for Peaceful Purposes was concluded in April 1976 between the United States and the Soviet Union. The PNET governs all nuclear explosions outside of the weapons test sites that were specified under the TTBT. The main provisions: not to carry out any individual nuclear explosions with a yield exceeding 150 kilotons; not to carry out any group explosion having an aggregate yield exceeding 1,500 kilotons; and not to carry out any group explosion having an aggregate yield exceeding 150 kilotons unless the individual explosions in the group could be identified and measured by agreed verification procedures.

1977: Trilateral CTB talks are resumed

Trilateral talks were resumed in Geneva between the United States, the United Kingdom, and the Soviet Union on a comprehensive test ban agreement.

1988: JVE

On December 9, 1987, the United States and the Soviet Union agree to conduct a joint verification experiment at each other's test sites (which occurred the following year).

1992: The United States declares a test moratorium

U.S. nuclear testing was halted by a unilateral moratorium codified in the Hatfield-Exon-Mitchell Amendment to the Fiscal Year 1993 energy appropriations bill. The amendment included a nine-month testing moratorium, restricted testing beyond the moratorium's timeframe, banned testing after September 1996 (unless other nations tested), and required a comprehensive test ban be negotiated by 1996. President George H. W. Bush (reluctantly) signed the law on October 2, 1992, though with reservations.[135] President Clinton extended this moratorium twice, and every U.S. administration after Clinton has also reaffirmed the nuclear testing moratorium.

135 von Hippel, Frank N., "The Decision to End U.S. Nuclear Testing," *Arms Control Today* 49, no. 10 (December 2019). https://sgs.princeton.edu/sites/default/files/2020-01/vonhippel-2019.pdf. Accessed September 16, 2020.

1993: Multilateral CTBT negotiations begin

On December 16, 1993, the United Nations General Assembly passes a resolution on supporting multilateral CTBT negotiations. The negotiations officially began in the Geneva-based Conference on Disarmament (CD) on January 25, 1994.

1995: NPT Review and Extension Conference

On May 11, 1995, the NPT state parties agreed to extend the NPT indefinitely. Most member states linked the issue of a comprehensive nuclear test ban to the NPT's extension, so the CTBT became part of the NPT "bargain." The conference adopted the Principles and Objectives for Nuclear Non-Proliferation and Disarmament, which specifically called for the conclusion of the CTBT negotiations by 1996.

1995: President Clinton endorses zero-yield

On August 11, 1995, President Clinton officially announced that the United States would pursue a zero-yield treaty and promised U.S. leadership in the global effort to conclude the agreement.[136]

1996: Conclusion of the CTBT negotiations

On September 10, 1996, the UN General Assembly adopted the Comprehensive Test Ban Treaty. On September 24, President Clinton was the first world leader to sign the CTBT. For Clinton, the CTBT was an important tool to support U.S. non-proliferation policy. He said that "a test ban can strengthen our efforts worldwide to halt the spread of nuclear technology in weapons."[137]

1999: Senate vote on the CTBT

The CTBT was submitted to the U.S. Senate in September 1997 for advice and consent. The Senate hearings on the CTBT began on October 5, 1999 in front of the Foreign Relations and the Armed Services Committees. On October 13, 1999, the Senate voted against the CTBT ratification by a vote of 51-48. Despite this outcome, the White House remained committed to the agreement. Even after the Senate rejection, President Clinton wanted to maintain momentum and appointed retired General John Shalikashvili, a former Chairman of the Joint Chiefs of Staff, to lead a task force and address Senate concerns about the CTBT.

2000: U.S. presidential campaign and the CTBT

During the 2000 presidential campaign, George W. Bush supported the extension of the testing moratorium, but opposed the CTBT. He argued that the treaty "does not stop proliferation, especially in renegade regimes. It is not verifiable. It is not

136 Clinton, William J., "Remarks Announcing a Comprehensive Nuclear Weapons Test Ban," The White House (August 11, 1995). https://www.govinfo.gov/content/pkg/WCPD-1995-08-14/pdf/WCPD-1995-08-14-Pg1432-2.pdf. Accessed September 16, 2020.

137 The White House, *Presidential Decision Directives: PDD/NSC-11–Moratorium on Nuclear Testing*, Office of the Press Secretary (July 3, 1993). https://fas.org/irp/offdocs/pdd11.htm. Accessed September 16, 2020.

enforceable. And it would stop us from ensuring the safety and reliability of our nation's deterrent, should the need arise."[138] Soon after his election, President Bush confirmed in his 2001 Nuclear Posture Review (NPR) that he was not interested in pursuing CTBT ratification, and his Press Secretary said that the White House did not rule out a return to testing in the future.[139]

2010: President Obama's Nuclear Posture Review (NPR) and the CTBT

The Obama administration took a 180° turn, declaring that CTBT ratification will be high on the arms control agenda. President Obama's 2010 NPR[140] argued that the CTBT was essential to strengthen the non-proliferation regime, and to gather international support for nuclear security efforts. The White House, however, did not submit the CTBT to the Senate for a vote as the New START ratification proved to be longer and more difficult than expected. Further, Democrats lost important seats in the Senate after the 2010 mid-term elections. Engagement on the issue continued but the political momentum was lost.[141]

2018: President Trump's NPR and the CTBT

In the 2018 NPR,[142] President Trump took the same position on the CTBT as the Bush administration. The document states that the White House will not pursue ratification of the CTBT, but it will continue the nuclear testing moratorium, unless the safety and effectiveness of the U.S. nuclear stockpile requires a return to nuclear testing.

2019: Calls to "un-sign" the CTBT

Technical arguments by a group of scientists and the growing intelligence concerns about potential Russian and Chinese non-compliance with the CTBT's zero-yield requirement fueled renewed advocacy against the CTBT and the testing moratorium. See the next appendix for a detailed discussion on the question of "un-signing" the treaty.

---

138 Bush, George W., "Presidential Election Forum: The Candidates on Arms Control," *Arms Control Today* (September 2000). https://www.armscontrol.org/act/2000-09/features/presidential-election-forum-candidates-arms-control. Accessed September 16, 2020.

139 Nikitin, Mary Beth D., C*omprehensive Nuclear-Test-Ban Treaty: Background and Current Developments* (Washington, DC: Congressional Research Service, RL33548, September 1, 2016). https://crsreports.congress.gov/product/pdf/RL/RL33548. Accessed September 16, 2020.

140 United States Department of Defense, *Nuclear Posture Review* (April 8, 2010). https://dod.defense.gov/Portals/1/features/defenseReviews/NPR/2010_Nuclear_Posture_Review_Report.pdf. Accessed September 16, 2020.

141 Interview with Rose E. Gottemoeller (February 19, 2020).

142 United States Department of Defense, *Nuclear Posture Review* (February 2, 2018). https://media.defense.gov/2018/Feb/02/2001872886/-1/-1/1/2018-NUCLEAR-POSTURE-REVIEW-FINAL-REPORT.PDF. Accessed September 16, 2020.

2020: White House discussions about nuclear testing

According to media reports,[143] senior Trump administration officials discussed in a "deputies meeting" on May 15, 2020 whether to reverse the 1992 moratorium and return to nuclear testing. In reaction to these reports, former Vice President Joe Biden stated that "The possibility that the Trump administration may resume nuclear explosive weapons testing in Nevada is as reckless as it is dangerous. We have not tested a device since 1992; we don't need to do so now…A resumption of testing is more likely to prompt other countries to resume militarily significant nuclear testing and undermine our nuclear nonproliferation goals."[144]

2020: The current status of the CTBT

As of July 2020, the CTBT has been signed and ratified by 168 states. From the 44 Annex 2 states, eight states have not ratified the CTBT. Among these eight holdout states, China, Egypt, Iran, Israel, and the United States have signed but not ratified the agreement, while the Democratic People's Republic of Korea, India, and Pakistan have not signed the CTBT.

143 Sonne, Paul, "Trump administration discussed conducting first U.S. nuclear test in decades," *Washington Post* (May 22, 2020). https://www.washingtonpost.com/national-security/trump-administration-discussed-conducting-first-us-nuclear-test-in-decades/2020/05/22/a805c904-9c5b-11ea-b60c-3be060a4f8e1_story.html. Accessed September 16, 2020.
Julian Borger, "US security officials 'considered return to nuclear testing' after 28-year hiatus," *The Guardian* (May 23, 2020). https://www.theguardian.com/world/2020/may/23/us-security-officials-considered-return-to-nuclear-testing-after-28-year-hiatus. Accessed September 16, 2020.

144 Quoted in: Arms Control Association, "Reaction to White House Nuclear Testing Proposal Strongly Negative," *ACA Issue Briefs* 12, no. 4 (June 16, 2020). https://www.armscontrol.org/issue-briefs/2020-06/reaction-white-house-nuclear-testing-proposal-strongly-negative. Accessed September 16, 2020.

Appendix 3

On "Un-Signing" the Treaty

As one indicator of continued deep opposition to the CTBT, a discussion has emerged within the policy community about how to remove from the legislative docket a treaty that has failed to receive the consent of the Senate.[145] In March 2019, four Republican Senators urged President Trump to "un-sign" the CTBT.[146] If a way were found to un-sign the treaty, this would oblige a future president committed to CTBT ratification to re-submit it to the Senate. It would not necessarily terminate the test moratorium.

The question of "un-signing" the CTBT raises a number of legal questions. According to Article 18 of the 1969 Vienna Convention on the Law of Treaties, once a state signs a treaty, it "is obligated to refrain from acts which would defeat the object and purpose of a treaty."[147] The United States signed the Vienna Convention in 1970, but, as with the CTBT itself, has not ratified it. However, the Vienna Convention is a codification of customary international law, has been recognized as such by the U.S. government, and thus it should be considered binding on the United States. Whether the CTBT itself has become customary international law is hotly debated.

Based on Article 18 of the Vienna Convention, the Clinton administration's decision to sign the CTBT means that the United States should refrain from acts that are against the object and purpose of the treaty. This implies that a U.S. return to nuclear testing would be a violation of a legal obligation. However, having been rejected for ratification there is now a question whether the rejection "cancels" the signing, and whether it means that the United States has no obligation to refrain from nuclear testing under international law.

Different administrations have had different interpretations of the obligations associated with a treaty that has failed to receive Senate consent to ratification. After the Senate rejection of the CTBT, President Clinton reaffirmed that his administration would continue the nuclear testing moratorium and the push for ratification. His administration concluded that the signing of the CTBT created a legal obligation for the United States, which was mostly evident in Secretary of State Madeleine

---

145 Spring, Baker and Dodge, Michaela, *Keeping Nuclear Testing on the Table: A National Security Imperative* (Heritage Foundation, February 27, 2013). https://www.heritage.org/arms-control/report/keeping-nuclear-testing-the-table-national-security-imperative. Accessed September 16, 2020.

146 Sonne, Paul, "U.S. military intelligence steps up accusation against Russia over nuclear testing," *The Washington Post* (June 13, 2019). https://www.washingtonpost.com/world/national-security/us-military-intelligence-steps-up-accusation-against-russia-over-nuclear-testing/2019/06/13/2dadf2e2-8e26-11e9-b162-8f6f41ec3c04_story.html. Accessed September 16, 2020.

147 United Nations, *Vienna Convention on the law of treaties* (May 23, 1969). https://treaties.un.org/doc/Publication/UNTS/Volume%201155/volume-1155-I-18232-English.pdf. Accessed September 16, 2020.

Albright's statement that "the United States will continue to act in accordance with its obligations as a signatory [of the CTBT] under international law..."[148]

The Bush administration had a different take on this question. In response to the request of Senator Kyl, Secretary of State Condoleezza Rice wrote a letter to clarify that "We do not believe the treaty imposes any current obligation on the United States resulting from U.S. signature in 1996, and we do not consider the United States to have obligations under international law as a signatory to the treaty."[149]

The Obama administration swung back to the Clinton perspective. Secretary of State John Kerry confirmed that "we are pursuing a political statement of the NPT's nuclear weapon states, all of whom are CTBT signatories, affirming their view that a nuclear test would defeat the object and purpose of the CTBT. As a matter of international law, treaty signatories are obliged to refrain from acts which would defeat the object and purpose of a treaty unless they make their intention clear not to become a party to the treaty."[150] This statement shows that the Obama team believed that signing the CTBT put the United States under legal obligation not to conduct nuclear tests, because the administration was seeking ratification.

The Trump administration has swung back to the view articulated by Secretary Rice.

148 Mendelsohn, Jack, "Still bound," *Bulletin of the Atomic Scientists* 56, no. 1 (2000), p42–43.

149 Rice, Condoleezza, "Secretary Rice's Response to Senator Kyl's Questions of July 17, 2006 on the CTBT" (September 7, 2016). Senate Hearing 114-724, p46-47. https://www.foreign.senate.gov/imo/media/doc/090716_Transcript_The%20Administration's%20Proposal%20for%20a%20UN%20Resolution.pdf. Accessed September 16, 2020.

150 Ibid. p39.

"International proposals to ban nuclear weapons and their testing began within six months of the first nuclear test at Trinity and months before the first peacetime nuclear tests at Bikini. Seventy-five years later, views remain highly polarized. There is no better example than the U.S. debate about ratification of the CTBT—which remains intense 25 years after the U.S. Senate's refusal to consent to ratification. Over the intervening period, both advocates and opponents of the CTBT have seen the strength of their arguments rise and fall, as well as new arguments come and go, in the context of a changing strategic environment, shifting arms control standards, growing concerns about Russian and Chinese compliance, and advancing technology. Anna Péczeli and Bruce Goodwin skillfully explore this complex landscape, with a primary focus on the technical issues in the ratification debate and subsequent lessons learned. In updating our understanding of the facts and issues through historical and technological analysis, Péczeli and Goodwin have provided a valuable service. And their timing is excellent. The quality of debate will be greatly enhanced if both proponents and critics of the CTBT read this book."

Ambassador Ronald Lehman

Former START I Negotiator and former Director, United States Arms Control and Disarmament Agency
Counselor to the Director, Lawrence Livermore National Laboratory

"A clear-eyed review of five technical issues that shape the debate over CTBT ratification. Especially well described are both sides of the argument about whether Russia and China are breaking the zero-yield barrier. The language will be easily understood even by non-experts, so this report should be required reading for anyone interested in nuclear testing."

Rose Gottemoeller

Former New START Negotiator and former Under Secretary of State for Arms Control and International Security
Payne Distinguished Lecturer, Freeman-Spogli Institute, Stanford University

Lawrence Livermore National Laboratory

ISBN-978-1-952565-05-2