DISCLAIMER

LLNL-TR-2013409

Stockpile Stewardship and Nuclear Testing: A Technical Assessment

G Miller

November 2025

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Stockpile Stewardship and Nuclear Testing

A Technical Assessment

Dr. George H. Miller

Summary

Since 1992, the United States has retained confidence in its nuclear weapon stockpile without performing any tests that produce nuclear yield. Instead, it has invested in a stockpile stewardship program (SSP) based on the same approach that was validated through fifty years of experience during the nuclear testing era. As the test moratorium continues and new information becomes available, it is both necessary and appropriate to periodically revisit the question of whether or not nuclear testing should resume. In a recent article by Dr. Mark Schneider[1], he asserts that the United States should return to nuclear testing to address the issues that negatively affect our country's nuclear deterrent. A serious, technically informed evaluation of the most significant of the issues Dr. Schneider has raised is provided in this paper. The conclusion of this evaluation is that the strategy of a well-funded SSP coupled with rigorous assessments to identify if there is a specific need for a nuclear test is sound and cost-effective approach. Any decision otherwise needs to take into careful consideration all the classified SSP successes and data pertaining to the particular issues in question.

Introduction

In a recent article, *The Case for Resumed Nuclear Testing*[2], Dr. Mark Schneider asserts that:

> "The United States faces a serious potential problem — the decline in the reliability of U.S. nuclear weapons resulting from more than 30 years without nuclear testing"[3]

Since late 1992, decision-makers have determined that the United States is best served by keeping its nuclear experiments sub-critical. Annual assessments continue to conclude that conducting slightly supercritical "hydronuclear" experiments or higher yield nuclear-explosive tests are not technically required and a return to nuclear testing would do more harm than good to U.S. interests. It is reasonable to periodically revisit this assessment to address advances in technical capability, the stockpile issues that need to be addressed, and changes in the geopolitical environment. While a complete and thorough evaluation requires access to classified information, much can be said in an unclassified environment.

This note will focus on the technical aspects of Dr. Schneider's three principal assertions:

- That science-based stockpile stewardship (SBSS), implemented through the Stockpile Stewardship Program (SSP), is inherently unable to provide adequate confidence in the country's nuclear deterrent:

  > " This is a seriously flawed concept. Virtually nothing is put into service without testing the finished product. When dealing with very expensive items like dams, bridges or buildings, testing is not to the point of destruction, but scale models are often tested to destruction and designs are normally very conservative with regard to strength. The zero-yield testing limitation excludes testing of functional scale models of nuclear weapons and one-point safety tests.

[1] Mark B. Schneider, *The Case for Resumed Nuclear Testing* (National Instituted for Public Policy, Occasional Paper, Volume 5, Number 9, September 2025)
[2] Ibid
[3] Ibid., pg. v

In most instances, SBSS is equivalent to putting new civilian or military aircraft into service without ever flight testing them."[4]

- That the current system of safeguards that includes the requirement for the directors of the three National Nuclear Security Administration (NNSA) laboratories and the Commander of the United States Strategic Command (USSTRATCOM) independently assess the need for nuclear testing is inadequate.

    "Under stockpile stewardship, the certification process has become a quasi-political exercise."

    "Vested interests have developed at the labs"[5]

- That nuclear testing is intrinsically less costly and more valuable than the Stockpile Stewardship Program.

    "Stockpile Stewardship is inherently costly and less effective than even minimal nuclear testing"

    "Life extension program should be relatively simple, prompt and inexpensive. Instead, due to the stockpile stewardship program, they are of long duration and expensive."

    "Even taking into account the startup costs of resumed nuclear testing, one year's cost of operating the laser fusion facility would assure the safety and effectiveness of the U.S. nuclear stockpile and even allow for the development of new and improved nuclear weapons in a fraction of the time and at much less expense than currently planned."[6]

To counter these assertions requires a fuller context for understanding the role that nuclear testing plays in assessing confidence in the U.S. nuclear weapons, particularly in view of the successes of the SSP and the process by which the Directors of the three nuclear weapons laboratories and the Commander of USSTRATCOM have assessed and provided high confidence in the safety, reliability and effectiveness of the nation's nuclear weapons stockpile for the past 30 years. Additional background is found in the article "Stockpile Stewardship: What were we thinking? How Did It Work Out"[7] which is liberally quoted in Dr. Schneider's article.

Background - Nuclear Testing in Context

Experimental data plays an important role in the development and maintenance of nuclear weapons, as with other technical endeavors. Unlike most engineered systems, however, nuclear weapons have never been tested in a statistically meaningful way — either for any single operational configuration or over the entire spectrum of operational conditions. From the very beginning, a different risk management approach has been used.

> The methodology involves extensive computational simulations from both a physics and engineering point-of-view, non-nuclear testing of both the physics and engineering aspects of the design, sufficient design margins, connection and comparison to nuclear test data, and extensive peer review.

[4] Ibid., pg.22. I would note that while all military and commercial aircraft undergo extensive test, almost all modern airplanes fly as designed the first time out.
[5] Ibid., pg. 57-58
[6] Ibid., pg. 49-50, 61
[7] George Miller, "Stockpile Stewardship: What Were We Thinking? How Did It Work Out?" in Brad Roberts, ed. *Stockpile Stewardship in an Era of Renewed Competition* (Livermore, CA: Center for Global Security Research, Lawrence Livermore National Laboratory, April 2022)

> An important part of the risk management approach has always been the existence of a robust surveillance and manufacturing infrastructure that allowed rapid recognition, assessment and replacement of problematic parts and systems.

This basic approach was augmented by the fact that the nuclear stockpile continuously evolved, allowing new understanding, new technologies and new concepts to replace older ones.

As noted in a published discussion of this issue:

> "With very few exceptions, no stockpiled weapon was ever tested in anything that resembles the way it was intended to be used. Compromises, in some cases extensive ones, were made. Technical judgement, computations etc., were used to infer the relationship of the actual tested device to the stockpile.
>
> "Which brings me to my most important perspective: 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 judgement of technical experts based on a rigorous process that considered all the available data, computational simulations, considerations of margins, etc."[8]

In addition to limitations on the number of nuclear tests that could be performed, the data that could be obtained from a nuclear test was very limited. The nature of a nuclear weapons is such that it is extremely difficult to obtain information about what is going on inside the device itself. In physics terms, a nuclear weapon is "optically thick" – meaning information from the inside of the device has a hard time getting to the outside of the device where it can be measured. This meant that integral quantities like yield could be inferred (although in a technical sense it too was never measured directly), as could some other gross features, but rarely were details unveiled that would tell the designer why it worked the way it did. The details had to be inferred from computational simulations augmented by non-nuclear experimental data.

The value of nuclear test data is also critically dependent on the yield of the test; the type of data available and its value is reduced as the yield is reduced. The technical implications of these thresholds were carefully considered in the discussions that led up to the U.S. decision to pursue a "zero yield" Comprehensive Test Ban Treaty in which only tests that never produce a super-critical assembly were allowed. So called "hydronuclear" tests, which are super-critical, were explored by the U.S. during the 1958–1961 moratorium. The nuclear yield in these tests produced no more than four-tenths of a pound of high explosive. Such tests are unreasonably costly with respect to the data they provide; other sub-critical experiments with use of high-quality diagnostics are significantly more important.

Over time, as more of the underlying science and technology became understood, as nuclear test diagnostics improved, as simulation capability and non-nuclear testing capability improved, the number of tests required to develop a new concept decreased significantly — as long as the design was reasonably represented by the past data base. This approach remains true today – the stockpile stewardship approach is valid as long as the modifications are well grounded in the past nuclear test data base and sufficient margins are present.

With the introduction of the "modern" nuclear weapons systems in the late 1960's, a number of important characteristics, excursions in terms material formulations and characterizations, engineering tolerances, etc. that are relevant to today's stockpile were nuclear tested to create a better understanding of the underlying limits of the technologies. The existing nuclear test data base provides a broad foundation for SBSS.

---

[8] George Miller, "Stockpile Stewardship: What Were We Thinking? How Did It Work Out?", op. cit., p 16.

Stockpile Stewardship Approach and Successes Provide Confidence in the United States' Nuclear Weapons

The assertion that reliability has declined without nuclear testing is based on a misunderstanding of the basis for making assessments. The basis for assessments of the stockpile has not fundamentally changed. Assessments are still based on expert judgement. **What has changed since the end of nuclear testing are both the questions that require attention and the tools available to provide input on the assessment.** Until very late in the nuclear testing era, the most serious questions were whether or not a new, untested concept or a significant extrapolation from a previously tested design would work as anticipated. There were a few confirmatory and "stockpile confidence" tests but as described earlier, these involved changes and compromises, often extensive, from the actual operational conditions. Today, the principal questions have to do with the effects of aging and changes in engineering and physics details associated with life extension programs (LEPs) of previously tested stockpile systems. These programs have so far produced weapons that are no more "new" than a refurbished 1980's Chevrolet. The nuclear weapons program is just beginning to enter a phase where designs are considered that differed in some details from previously tested designs but remain in a well-tested "design space." The designs under consideration maintain similar military characteristic (size, weight, yield, etc.); however, their physical instantiation could be considered "new." In these cases, the questions being asked will be intermediate between those asked of a completely new untested concept and detailed questions about the effects of engineering and aging changes.

Thirty years of investments in SBSS have vastly increased knowledge about nuclear weapon performance. **Modern SSP tools for nonnuclear testing of components and materials along with past nuclear testing data provide the data to improve and validate the performance of computer simulations**. **It is this suite of capabilities that provide the foundation for the expert judgement necessary to sustain confidence in the stockpile.** High-energy-density science facilities, such as the National Ignition Facility (NIF) have enabled scientists to improve the modeling of materials—including special nuclear materials—at the temperatures and pressures of a nuclear explosion and are not attainable from diagnostics on nuclear explosive tests. Radiographic capabilities enable detailed examination of imploding pits. The capability of supercomputers has improved by a factor of 10^7 over the past thirty years. Today's computational simulations include physics that was known to be important and unable to be accommodated by the computers of the past, along with the "real-world" engineering associated with actually manufacturing and attaching a nuclear device to its military system. With the combination of vastly improved data from nonnuclear tests and validated high-resolution 3D simulation codes, stockpile stewards have been able to resolve some unanswered questions about performance that persisted when nuclear testing ceased by replacing adjustable parameters with data-based physics models. The use of these tools and the knowledge gained has enabled a much deeper understanding of nuclear weapon performance than ever before, and their use has sustained confidence in the safety, reliability and effectiveness of the stockpile.

The data from experiments are key to understanding important effects like material aging and manufacturing tolerances. In the case of some of the major issues like aging, changes in the detailed materials properties, like material strength, crystalline state, inclusions are important because they have the potential to affect performance. **Fortunately, these properties and the underlying morphology can be measured today, and experiments to evaluate the effect of changes on important nuclear weapon characteristics can be performed. The experimental data also feed part of the validation process for the computational simulations.**

The capability of today's computational simulations is significantly enhanced over that used to design nuclear weapons during the testing era. Knowledge of material properties in weapons-relevant regimes has also greatly improved. The validity of simulation capability is quantified by both comparison against detailed non-nuclear data and against the more integral measurements of past nuclear tests. Far from the "historical research" referred to by Dr. Schneider,[9] **using the vast reservoir of data from past nuclear tests to validate the performance of modern simulations is essential to providing a basis for confidence in the assessment of current issues.** In the era of nuclear testing, extensive calibration of the simulations to individual tests were required to provide some ability to extrapolate to conditions different from that particular test. In fact, different calibrations were often required for different tests within the same design family. Today's simulations are much more capable and accurate and are simultaneously validated against all the tests of broad categories of warhead design, often without significant calibration changes being required.

Equally important, these same simulations are used to evaluate many of the issues referenced in the 1987 report on the role of nuclear testing in ascertaining the reliability of the nuclear stockpile.[10] Some of this data was also discussed by Dr. Schneider in his article.[11] While much of this data set refers to issues that occurred early in the development of the modern stockpile and is extremely unlikely to appear in today's considerations, this data, and in fact the whole data set of U. S. nuclear tests is extremely valuable in ascertaining the validity of modern computational simulations. **Many of the previous "anomalies" have now been explained and significant computational normalizations have been eliminated**. There is of course, more to do in this area as it remains beyond the state of the art to describe the entire data base with a single approach to simulation. The program is rapidly moving from being able to match a single warhead design's test history to being able to match broad categories of devices, for instance all those with insensitive high explosive.

Unlike the assertion that "the ability to make such predictions of the nuclear performance of weapons in their current state has not been demonstrated, and cannot be demonstrated, without a nuclear test program,"[12] the current program does this on a regular basis. While it is certainly true that none of the stockpiled weapons have been nuclear tested in the current aged state, the program addresses these issues in different ways. In some cases, for example the aging of high explosives, the effects can be directly measured. In other cases, the effects can be measured in scaled experiments and the computation simulations validated because the relevant physics equations are manifestly scalable. In some situations, it is the underlying material morphology that has changed and the effects of these changes can be measured and compared to unaged samples. In all cases, the uncertainties are carefully evaluated against the margins to failure so that an objective picture of the risks can be evaluated. In all cases to date, the evaluated risks have been minimal, similar in many respects to the compromises made to conduct a "stockpile confidence" test. The judgement has been made that a nuclear test would not substantially reduce the already minimal risk.

Another important part of the approach to stewardship of the nuclear weapons stockpile is to increase the margins against failure at every opportunity. **Unlike the assertions made by Dr. Schneider that increasing robustness is precluded because increasing robustness requires adding fissile material that is precluded by on-point safety concerns[13], there are other methods to increase robustness (margin against failure) that do not require the addition of fissile material or introduce one-point safety concerns. These methods have been applied whenever needed.**

---

[9] Mark B. Schneider, *The Case for Resumed Nuclear Testing,* op. cit. pg. x

[10] George H. Miller, Paul S. Brown Carol T. Alonso, *Report to Congress on Stockpile Reliability, Weapons Remanufacture and the Role of Nuclear Testing*, Livermore, CA; Lawrence Livermore National Laboratory, October 1987)

[11] Mark B. Schneider, *The Case for Resumed Nuclear Testing,* op. cit. pg.6 — 22

[12] Ibid., pg. 23

[13] Ibid., pg. 46

Most of the successes of stockpile stewardship remain classified. On several occasions however, unclassified examples have been published by the laboratories.[14] In one of the articles Brad Wallin, principal associate director for nuclear weapons program at LLNL, named six areas of study that exemplify the scientific discoveries important to stockpile stewardship over the past 30 years.

> "**Some of the most important areas have included the precision with which we've been able to understand the behavior of plutonium; solving the energy balance problem; discoveries made at the National Ignition Facility; detailed understanding of the chemical behavior of high explosives; advanced manufacturing techniques and new materials; high physics fidelity, and high-resolution simulation codes**."

One of the most notable accomplishments of stockpile stewardship was solving a mystery that had confounded some of the smartest physicists for five decades: the mystery of missing energy produced during nuclear tests. The solution involved theoretical advances, improvement in simulation capability, and experimental validation.

> "The energy balance problem was first recognized in the 1960s when Livermore developed the first two-dimensional radiation hydrodynamic simulation tools. Over many decades when those tools were applied to conducted nuclear tests, it appeared that the tests violated a basic principle of physics known to every college freshman physics student: conservation of energy."

> "When underground testing ended in 1992, solving the energy balance mystery became important because weapons designers could no longer rely on new test data to validate their codes."

Perhaps the most widely recognized program success, whose implications for stockpile stewardship have been barely noted, is the achievement of the ignition of a thermonuclear fuel (fusion) at the National Ignition Facility in December 2022 and the continued successes in improving the behavior of an ignited capsule. **The implications for stockpile stewardship demonstrating ignition on NIF are significant and vital for sustaining confidence in and modernizing the stockpile.**

- While the National Academy of Sciences report in 1997 stated that "the achievement of ignition appears likely", privately they gave it a 50/50 chance[15]. Achieving ignition at NIF, like stockpile stewardship, was a grand challenge successfully executed. The first burning plasma experiment used slightly more than 1.8 MJ (1.93 MJ), and subsequent higher yield shots have used over 2 MJ and produced greater energy gain. The success at NIF built on the same revolutionary improvements in computational simulation, experimental diagnostics, and materials science as those required by stockpile stewardship. Both efforts are founded on a similar culture that is at the same time bold, cautious and precise in implementation, and willing to take well informed risks.
- NIF's success provides confidence that the stockpile stewardship approach to the problems inherent in multi-disciplinary science and engineering, with complicated non-linear physics, significant materials and engineering challenges can be solved with the scientific and technological revolutions that are taking place.
- In addition to the successful demonstration of ignition, NIF has provided an essential experimental platform to directly address a variety of stockpile issues and provide experimental validation that changes occurring due to aging or engineering differences do not affect the overall confidence in the nuclear stockpile.

[14] *Stockpile Stewardship at 20 years*, in Lawrence Livermore National Laboratory Science and Technology Review, July/August 2015 and *Scientific Discovery for Stockpile Stewardship*, Lawrence Livermore National Laboratory, LLNL-MI-840097, 27 September 2022.
[15] Private communication from Dr. John Lindl.

- NIF also provides a magnet for highly qualified and talented individuals and a training ground for both how to solve complicated, multidisciplinary challenges, and importantly, create a critical culture of constant questioning of assumptions and humble realism.

In a very real sense, the weapons program has a significantly more capable and a deeper understanding of the detailed nature of nuclear weapons, their sensitivities, uncertainties, and risks as a result of stockpile stewardship than was ever possible in the nuclear testing era. These capabilities and understandings continue to allow the careful evaluation of risks and underwrite the assessment of confidence in the proper safety, security and effectiveness of the stockpile and whether a nuclear test is required to retain adequate confidence in the deterrent. Equally important, these capabilities allow informed assessments and decisions to be made about how extensive modifications of previously tested designs can be made with acceptable risk.

The Annual Assessment Process is Robust

When the U.S. decided to forgo nuclear tests and pursue a zero-yield test ban, an annual, formal rigorous process was created to directly focus on assessing the health of the U.S. nuclear weapons and ensuring high-level confidence in the stockpile. The process specifically calls for recommendations on whether or not a nuclear test was required to ensure adequate confidence in the nuclear deterrent. This process required that the directors of the three NNSA nuclear weapons laboratories and the Commander of USSTRATCOM independently provide their assessment to the President of the United States through the Secretaries of Energy and Defense. In this process it was recognized that while there may be multiple assessments about the health of the stockpile, there is only one set of data. This data is annually presented for review by all individuals and organizations with nuclear weapons responsibilities.

The responsible design teams present their evaluations for an extensive and critical review by the Directors. In addition, the Directors receive and independent assessment of the systems they are responsible for by teams from the other laboratory as well as independent internal "red teams." **These independent reviews in my experience are true "peer-reviews" in the sense that they are performed by true peers who have independently done work and provided alternate points-of-view about the data, experiments, and calculations that are relevant.** Also in my experience, every step in the process was extremely intense, with all the assumptions and uncertainties examined.

To provide input to Commander of USSTRATCOM, another completely independent panel of experts in nuclear weapons evaluates the data presented and provides their views on the issues and risks that arise and whether or not a nuclear test is required to resolve the issues or to reduce the risks to tolerable levels. The Commander of USSTRATCOM then evaluates data about the warheads in the context of his mission responsibilities.

Just like the situation during the nuclear testing era, these assessments and attendant statements of confidence are not a proof. They are, and have always been, the best judgement of experienced, highly capable experts, high integrity scientists, technologists, and engineers. Confidence in the stockpile rests on the credibility of the individuals and the integrity of the processes that have been put in place.

Nuclear testing would be an expensive addition to the current program

The cost of returning to a nuclear testing-based program is significantly more expensive than estimated by Dr. Schneider. For instance, he asserts that a nuclear test program could be executed for the cost of the annual operating budget for the National Ignition Facility. **The annual operating budget of NIF is approximately $400M. In 1985 the annual nuclear testing budget was $548M, which is $1.649B in today's dollars.** Given the extreme escalation in the cost of doing business for non-inflationary reasons, it is likely that the cost of returning to nuclear testing is substantially in excess of this constant dollar evaluation. Many studies of the cost of resuming a nuclear test program point out the many rules and regulations that have changed and support the assessment that returning to nuclear testing is an expensive proposition.

More important is the opportunity cost of returning to nuclear testing — i.e. what would you have to give up. The entire Research, Testing and Engineering budget of the National Nuclear Security Administration is only $3.280B. A return to nuclear testing would decimate the existing program of assessment and evaluation and is a very poor trade-off.

Dr. Schneider points out, the cost of the nuclear enterprise and timescales for accomplishing work are excessive. This malady afflicts virtually every part of the defense establishment. For the nuclear enterprise, the lack of nuclear testing is not the issue. The major bottleneck is in fact the capability and capacity of the production complex. Careful analysis concludes that Dr. Schneider's assertion that the lack of nuclear testing is a major or even contributing factor does not hold water. A brief personal experience serves to illustrate the point.

> In 1974 I was asked to join the B77 secondary design group and assist with the design effort that was underway. The design under consideration involved some new concepts for which the available design tools did not yet exist. It did not perform as anticipated. I was asked to take over leadership of the design group. A completely new design was developed, manufactured, and tested at nearly full yield in approximately 4 months. It worked as designed.

The very short timescales were the result of:

- An overarching recognition of the importance of the mission at every level of organization
- A true partnership between the federal government, the laboratory and the production complex
- Judicious risk taking
- Alignment of authority and responsibility
- Appropriate processes
- A substantial, capable infrastructure
- Availability of appropriate design tools (given the risks that were being taken)

A drift away from these organizational concepts and a culture that practices them are the principal reasons for the current fragility of the nuclear enterprise — not anything associated with the inability to perform a nuclear test.

Assessment of Risks in the absence of nuclear testing

Like almost all areas of national security, confidence in the nuclear weapons that forms an important part of the country's deterrence posture is an exercise in risk management. **Because nuclear weapons have never been statistically or fully tested, there were risks when the U.S. could execute nuclear tests just as there are now when we cannot.**

Quoting again from, "Stockpile Stewardship: What Were We Thinking? How Did It Work Out?" three major risks remain today[16]:

- **The production complex.** The current production complex is fragile at best, unable to respond quickly to problems that might occur in the stockpile or the need for new capabilities in response to adversary actions. The timescales for any response are long, it remains very expensive and lacks the flexibility or capacity to handle multiple problems at once.

[16] George Miller, "Stockpile Stewardship: What Were We Thinking? How Did It Work Out?", Ibid., p 16.

- **The warheads.** The deployed warheads are design legacies of the Cold War. The United States is just beginning to implement designs that, while functionally and militarily the same, are more robust and easier to manufacture and maintain. While the ability of the nuclear enterprise to assess and extend the life of the deployed stockpile has been very successful, exercising the skills necessary to make major modifications has atrophied, and there are risks and benefits associated with this task.
- **The laboratory complex.** While more experimental and computational capability would be helpful, the major risk for the laboratory complex is overconfidence. Stated most succinctly: Do the people making the judgements about the stockpile have an appropriate understanding for what they know and what they don't know?

As discussed above, nuclear testing is neither the cause nor the solution to the issues afflicting the production complex or implementing appropriate changes to the deployed stockpile. The question of overconfidence is more nuanced. A few "demonstration" tests, as suggested by Dr. Schneider, would do little in my view to enhance the humbleness and careful attention to detail required of a nuclear weapon designer. Since the tests are likely to work as planned, they could in fact add to a sense of overconfidence. An extensive and robust test program as was executed in the past, with attempts at new and difficult designs, the allowance of failure could certainly deal effectively with overconfidence. In my judgement it is neither necessary or cost effective in the current circumstances.

Recommendation

Much has changed since the inception of the Science Based Stewardship Program that began more than 30 years ago. In addition, much of the data referred to in arguments in favor of a return to nuclear testing is unclassified and from the early days of the evolution of the country's understanding of detailed physics and engineering of nuclear weapons.[17] While the process for the Annual Assessment of the stockpile is extensive and robust in my view, maintaining confidence in the stockpile is a very important issue. Thus, having a mandated, extensive external review of the basis for confidence that takes into account all of the data, much of which remains classified, seems entirely appropriate. At a minimum, such a review should evaluate the three issues I have identified as areas of continuing risk as well as the relationship of any of the evaluated risks to nuclear testing. The most recent such study was done in 1999.[18]

Conclusion

Dr. Schneider's paper makes three important assertions

- SSBS is inherently unable to provide confidence. An evaluation of the technical data concludes:

  The tools of stockpile stewardship have provided physics and engineering data used to vastly improve and validate detailed models of performance on computers that are 10^7 times more powerful than 30 years ago.

  The tools have enabled stockpile stewards to better understand weapon performance, and they have used the improved models to both (1) explain unknowns about weapon performance when testing halted and (2) use the knowledge gained to ensure that designs have a high margin of conservatism to ensure performance of the "physics package" is far from the possibility of failure.

---

[17] For example, *The Consequences of a Comprehensive Test Ban Treaty*, Report of Senator Dewey F. Bartlett to the Committee on Armed Services, United States Senate, August 11, 1978, and George H. Miller, Paul S. Brown Carol T. Alonso, *Report to Congress on Stockpile Reliability, Weapons Remanufacture and the Role of Nuclear Testing*, Livermore, CA, October 1987

[18] *FY 1999 Report of the Panel to Assess the Reliability, Safety and Security of the United States Nuclear Stockpile*, Harold M. Agnew, John S. Foster, Sydell P. Gold, Stephen J. Guidice, James R. Schlesinger; November 8, 1999

Other nonnuclear modes of failure—and concerns about materials aging—are testable with capabilities that are vastly superior to technology 30 years ago.

- That annual certification is inadequate. An evaluation of the technical data concludes:

  SSBS includes a rigorous stockpile surveillance program and a dedicated effort to understand the aging of materials and impact on performance through a combination of modeling and experiments.

  The annual assessment process draws on surveillance efforts and expert judgment to evaluate and address arising concerns—and includes rigorous red teaming.

  Annual assessments serve the function of ascertaining if a test is necessary, what that test should be and what data needs to be gathered to restore confidence. Conducting an expensive test should be targeted at a specific need and that need can only be identified through a rigorous assessment program founded on the full range of SSBS capabilities.

- That the knowledge gained in a nuclear test is intrinsically worth the investment. An evaluation of the technical data concludes:

  Expert judgement is the foundation for confidence and nuclear tests were only one component—both expensive and without providing fundamental data about performance that SSBS data now provides.

  Nuclear testing never gave statistically significant information about reliability and invariably included some modifications to the designed weapons and could only be tested at one set of conditions the weapon might encounter.

  The annual assessment process draws surveillance efforts and expert judgment to evaluate and address arising concerns—and includes rigorous red teaming. An expensive and politically treacherous decision to conduct a test should serve a specific purpose.

  A high degree of attention is being paid in modernization programs and new designs to increase "nuclear package" performance margins to further reduce the value added of a final development nuclear test.

All national security activities involve management of risk. The nuclear weapons enterprise in the nuclear testing era and today are no different. Significant progress has been made toward implementing the vision of the stockpile stewardship program that was set more than 30 years ago. That trend continues. There are certainly important risks today that require immediate and urgent attention. The lack of nuclear testing to certify changes that are occurring in the existing stockpile due to aging or to certify new, more robust warheads is not among the major concerns or risks.

More than 30 years ago I concluded that forgoing nuclear testing involved significant risk:

> "We are not ready today for significantly reduced nuclear test limits. Until we can find ways to meet our responsibilities for ensuring the reliability and effectiveness of the U.S. nuclear weapons and ways to prevent the erosion of nuclear weapon expertise and judgement under restrictive nuclear test limits, it would be imprudent to commit this country to a regime of further nuclear test limitations."[19]

Today, I conclude that nuclear testing is not currently required to maintain high confidence in the safety, security and effectiveness of the nuclear warheads that provide part of the country's deterrent posture. This evaluation is based on my personal observation of the judgements made by today's stockpile stewards using the extensive data from past nuclear tests and the significantly enhanced computational and experimental

[19] George H. Miller, testimony in front of the House Armed Services Committee, Panel on the Department of Energy Defense Nuclear Facilities, March 31, 1992.

capabilities put in place by the Stockpile Stewardship Program. I judge that we have found ways “to meet our responsibilities for ensuring the reliability and effectiveness of the U.S. nuclear weapons.”

At the time the U.S. ceased nuclear testing, the NNSA lab directors acknowledged that a science-based approach to stockpile stewardship was a calculated risk and success was not guaranteed. In that context, they asked for a specific set of safeguards including a rigorous annual assessment process. The SSP has been phenomenally successful because of the investments the nation has made in experimental and computational capabilities to study the initiation of nuclear weapons and the underlying physics of weapon performance. These investments have served to drive U.S. leadership in many areas; this is particularly evident in supercomputing. No other country has similar capabilities. Coupled with the extensive data base from past nuclear tests, the U.S. is in a position of strength, not weakness. The knowledge gained has underpinned our ability to sustain and modernize the stockpile without nuclear explosive testing. This success is a demonstration of international leadership, technical prowess, and our confidence in the expert judgement of the nation’s stockpile stewards. These stewards will identify the need for a nuclear test should need arise. While there remain important risks to the nuclear weapons enterprise that must be mitigated, a return to nuclear testing is not the solution to any of them.

The strategy of a well-funded SSP coupled with rigorous assessments to identify if there is a specific need for a nuclear test is a sound and cost-effective approach. Any decision otherwise needs to take into careful consideration all the classified SSP successes and data pertaining to the particular issues in question.

The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC.

About the Author

Dr. George H. Miller is a Director Emeritus of Lawrence Livermore National Laboratory (LLNL). From 1972 to 1980 he was a member of one the nuclear weapons design divisions at LLNL and was responsible for more than a dozen nuclear tests, the development of the B77/B83 and W84 thermonuclear stages and was program manager for the development of the W84. From 1980 through 2006 he held a variety of management positions including head of the LLNL nuclear weapons program and the program responsible for construction of the National Ignition Facility. He was a major participant in the development of the Stockpile Stewardship Program. He became LLNL's tenth Director in 2006 and retired in 2012.