JOHN C. HOPKINS AND DAVID H. SHARP

THE SCIENTIFIC FOUNDATION FOR ASSESSING THE NUCLEAR PERFORMANCE OF WEAPONS IN THE US STOCKPILE IS ERODING

Current worldwide geopolitical developments—including questionable Russian pronouncements pertaining to nuclear weapons, concerns about nuclear proliferation in a volatile Middle East, tension between the strategic interests of nuclear-capable China, and the security interests of the United States' allies in the Pacific and the aggressive nuclear weapons program that is under way in North Korea—reaffirm the importance of effective nuclear deterrence now and for the foreseeable future.

The effectiveness of the US nuclear deterrent is based in part on an assurance, or "certification," that weapons in the stockpile will function as intended. Specifically, certification is a statement that assessments of the performance, safety, and reliability of the nuclear weapons allow the United States to conclude, with high confidence, that these weapons will function as expected. Thus, far from being a mere technical term, certification embodies a process that, along with scientific judgment, uses every bit of available evidence, especially that from applicable nuclear tests, to decide whether weapons of a particular type can be added to, or remain in, the stockpile. The concern that motivates this article is that the scientific foundation for assessments of the nuclear performance of US weapons is eroding as a result of the moratorium on nuclear testing.

What are the specific causes of our concern? First is the fact that the physical state of weapons in today's stockpile differs from what it was when their nuclear performance (e.g., yield) was tested, and second, the current nuclear test moratorium precludes a decisive determination of whether these changes in physical state adversely affect performance. Changes can result from aging, measures employed in life extension programs to mitigate the effects of aging, and deliberate modifications (e.g., to enhance manufacturability), and may even include unknown changes from unknown sources.

We emphasize that the above concern is not allayed by the widely recognized progress made by the Stockpile Stewardship Program (SSP) in advancing the scientific understanding of the operation of nuclear weapons. The importance of the results obtained by the SSP is not at issue here. Nevertheless, impressive though these results are, it has not been demonstrated that SSP-based results are, or will be, sufficient to supplant nuclear tests as a source of information that is indispensable for assessing the nuclear performance of the weapons in today's stockpile in a credible and trustworthy manner.

Most people acknowledge that confidence in nuclear performance is higher when one has nuclear test data. That is a "no-brainer," people like to say. Well, yes. But, some people ask, are nuclear test data all that important? Can't one live without them? What does the absence of nuclear test data really cost in terms of confidence in the performance of the US nuclear stockpile? The answer, apparently not a no-brainer, is that it may cost a lot. Here's why:

- Nuclear tests gave decisive, direct evidence about the behavior of new weapons destined for the stockpile. Each type of weapon currently in the stockpile was subjected to dedicated nuclear tests. Virtually no comparable data exist on the nuclear performance of stockpiled weapons in their current state.
- Scientific judgment is an important part of all performance assessments. For example, it was always necessary to extend assessments of a weapon's nuclear performance from a limited set of tested conditions to the full set of potential deployment conditions. Such extensions are ultimately a matter of scientific judgment. Nuclear testing provided a solid foundation for the development and evaluation of scientific judgment because it unequivocally tested performance predictions. Ironically, a no-test paradigm places

increased demands on scientific judgment, while removing the foundation for that judgment.
- A nuclear test moratorium leaves important questions unanswered. To see why, recall that in the absence of nuclear test data for weapons in their current state, predictions of their performance can draw on comparisons with historical nuclear test data, current laboratory experiments that stop short of explosive nuclear reactions, and simulations—but not on information from new nuclear tests.

When the nuclear weapons currently in the stockpile were designed and manufactured, there was no expectation that they would last indefinitely. The federal government's nuclear weapons laboratories have a continuing responsibility for maintaining the safety, performance, and reliability of these weapons as long as they remain in service. To fulfill this responsibility, one must be able to identify any serious performance problem that is present, fix the problem if possible, and determine whether the weapon is still certifiable.

This responsibility cannot adequately be met within the SSP. The SSP does allow a nuclear test to be requested if necessary to resolve a serious problem affecting a nuclear weapon. But this approach does not take into account the fact that in the past some serious problems were revealed only as a result of a nuclear test or that nuclear test data are required to develop and validate the scientific judgment and computer codes that must be used to assess the nuclear performance of weapons.

Such assessments of weapons in their current state are needed now, not sometime in the indefinite future, because the nation is carrying out life-extension programs and other modifications of the stockpile today. In the absence of experimental data that comes uniquely from nuclear tests, the nation can be left not knowing whether there are nuclear performance problems with any of its weapons that have not been revealed by surveillance programs. Likewise, it may not be known which problems that may come to light, by whatever means, could have a serious impact on the nuclear performance of weapons in the stockpile.

We note further that in the absence of testing the current stockpile maintenance program inevitably promotes a shift in the standard of assurance for the performance of stockpiled weapons from "decisive, direct evidence of proper performance" to "absence of evidence of unacceptable performance." Again, here's why:

- The ability to design nuclear weapons is a skill that is needed to support modernization and remanufacture of the evolving nuclear arsenal in a way that is truly responsive to current national security requirements. But the nation must confirm that the nuclear weapons program has retained this skill by successful application of the design process—design, build, and, of course, test.
- Some observers would argue that departures from a tested configuration (of a legacy weapon) due to aging or other factors can be expected to be "small," so that they are unlikely to have a significant effect on nuclear performance. Perhaps so, but the correctness of this assertion needs to be demonstrated rather than assumed. There is no guarantee that small changes in the state of a weapon will have small effects on its performance, owing to strong nonlinearities in the system dynamics and cumulative and cooperative effects arising from various sources.
- Confidence that today's nuclear weapons will perform properly is predicated on the assumption that there will be no surprises; that is, no significant departures from expected weapons behavior. How can scientists or policy-makers be so sure? The history of testing complex systems, nuclear and nonnuclear, is punctuated by unpleasant surprises. Although present-day designers are very bright and very careful, the same could be said about the scientists and engineers of the Cold-War period, when there were enough surprises to keep one's confidence in check.
- It is important to note that legacy warhead designs often had to be highly optimized so as to achieve the required yield while satisfying tight limits on the weight and size of the delivery system and rigorous safety and security requirements. In highly optimized designs, small defects can seriously impair performance. The result is that some small details (but which ones?) must be accounted for in making predictions. For nuclear weapons this can be a very difficult undertaking, and its eventual success in the absence of relevant data cannot be assured.

The above arguments are not ones that proponents of a continuing nuclear test moratorium or a Comprehensive Test Ban Treaty wish to hear. They have, naturally, argued that alternative methods can supply the decisive, direct evidence about nuclear performance that is needed to maintain confidence in the stockpile. Instead of ignoring or dismissing these arguments, we will briefly discuss some of them here.

For example, it is sometimes claimed that modern computer simulations can provide the basis for accurate and reliable assessments of nuclear performance, so that vital information formerly obtained from nuclear tests is no longer needed. But the claim is not correct, as the following analysis shows.

Nuclear performance simulations have historically played an important role in the design of nuclear weapons. During the nuclear testing era there were generally significant discrepancies between the output of these simulations and the nuclear test data with which they were compared. The parameters and physical inputs used in the simulations were commonly adjusted to reduce these discrepancies, using a procedure called calibration. This produced a modified simulation that was accurate enough that it could be used with reasonable confidence in an interpolative mode to explore numerous design options before settling on a final design.

One of the approaches being used today to develop nuclear performance simulations continues to rely on calibration, but it is transitioning to calibration to data from carefully designed above-ground nonnuclear experiments (AGEX). However, the simulation that is calibrated using AGEX data is not accepted for use unless the results agree with nuclear test data taken decades ago for the device being assessed today.

But therein lies the problem. The vast majority of archival nuclear test data pertain to newly manufactured devices. There were no aging or refurbishment effects present. Without nuclear testing, it is not possible to know which, if any, of the devices that were previously tested are representative of devices currently in the stockpile. Current performance predictions thus rely on simulations, and scientific judgment in the use of those simulations, that have never undergone an actual test to see how well they work. Considerable effort and resources are being directed toward developing nuclear performance simulations that will have comprehensive predictive capabilities, meaning that the results of an experiment can be reliably predicted using only information that is available prior to that experiment.

It is certainly conceivable that simulations with the requisite predictive capabilities could be developed. It is also true that weapons simulations have been improved. So, it is not our purpose to argue that predictive simulations for nuclear performance cannot be developed. Maybe they can be, but when?

The real question is how scientists would know whether such a predictive capability had actually been achieved. Certainly, careful science, including simulation verification and validation using the results of related nonnuclear experiments, could provide some confidence in a proposed predictive capability. But confidence gained in this way is not sufficient to certify nuclear performance simulations as replacements for nuclear tests. Why not? Because the only unequivocal way to demonstrate that predictions made with simulations meet expected standards of confidence is by establishing a track record of correct and reliable predictions that have been made using that simulation. However, 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.

It is also argued that aging of nuclear weapons can be circumvented by remanufacturing them. This is true, but the United States cannot implement this option today. At this time, it can remanufacture weapons in only very small numbers. Moreover, the nation is not prepared to conduct nuclear tests expeditiously to provide reality checks on the fidelity of the remanufacturing process.

A third argument is that since the inception of the SSP 25 years ago, no problems with existing nuclear weapons that must be resolved with a nuclear test have come to light. However, this point overlooks three important facts. First, the argument does not deal with the vital role that past nuclear tests have played in revealing serious nuclear performance problems and in validating nuclear performance simulations. Second, the argument ignores the possibility of unpleasant surprises—unknown unknowns. Finally, the surveillance program has had its ups and downs through the years, and has not always enjoyed a high priority, as pointed out by the Government Accountability Office.

Another contention is that if surveillance or inspection of a nuclear weapon reveals anomalies, that weapon can simply be removed from the stockpile. But this strategy could within a short time result in a nuclear stockpile whose composition and capability are determined not by design but by deterioration, obsolescence, or other unanticipated factors.

This approach is faced with other difficulties as well. Simply identifying anomalous features associated with aging or remanufacture of a nuclear weapon is not the whole story because surveillance alone cannot reveal what the effect of such anomalies on nuclear performance is going to be. In the absence of nuclear tests, it is necessary to rely on simulations to predict their effects in sufficient detail for stockpile stewardship. But the simulations have not been validated for this purpose.

Confidence that complex systems will function as intended must ultimately rely on proven means for conducting reality checks that will, with high probability, bring to light improper system performance. "Uncontrolled risk" is inherent in systems that lack such reality checks. In the absence of nuclear testing, the nuclear weapons program is exposed to the uncontrolled risk that assessments are significantly in error.

Not much can be done about this in the absence of experimental data on the nuclear performance of today's weapons. There are no easy fixes, and those that are feasible may have long lead times. Those who dismiss the need for nuclear test data pertaining to weapons in their current state are gambling with the nation's nuclear deterrent. Is this a good bet?

John C. Hopkins *is a nuclear physicist and a former leader of the Los Alamos National Laboratory's nuclear weapons program.* ***David H. Sharp*** *is a former laboratory fellow (retired) and a guest scientist at Los Alamos National Laboratory. He has served as leader of the Complex Systems Group and as chief scientist in the Science, Technology, and Engineering Directorate at Los Alamos. This document has been assigned report number LA-UR-18-23392 and approved for public release by the Los Alamos National Laboratory.*