
**The Nonproliferation Review**
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t716100717

# WHEN DOES A STATE BECOME A "NUCLEAR WEAPON STATE"?

Jacques E. C. Hymans

Online publication date: 18 January 2010

**To cite this Article** Hymans, Jacques E. C.(2010) 'WHEN DOES A STATE BECOME A "NUCLEAR WEAPON STATE"?', The Nonproliferation Review, 17: 1, 161 — 180
**To link to this Article: DOI:** 10.1080/10736700903484728
**URL:** http://dx.doi.org/10.1080/10736700903484728

# WHEN DOES A STATE BECOME A "NUCLEAR WEAPON STATE"? An Exercise in Measurement Validation

**Jacques E.C. Hymans**

*When does a state become a "nuclear weapon state"? How we choose to answer this question has significant implications for proliferation assessment, analysis, and policy. Traditionally, the standard demarcation line has been a state's first nuclear test, but in recent years analysts have increasingly focused instead on the accumulation of a significant quantity (SQ) of fissile material. The article argues that although the test/no-test indicator clearly has problems, its replacement by the SQ/no-SQ indicator would be highly counterproductive. The article instead proposes supplementing the traditional test/no-test indicator with a theory-driven approach that focuses on the incentives and disincentives to test.*

KEYWORDS: Nuclear weapon states; non-nuclear weapon states; nuclear testing; fissile material



When does a state become a "nuclear weapon state"? Traditionally, the standard demarcation line has been a state's first nuclear test. Before the test, states are considered "non-nuclear"; after it, they are considered to have "gone nuclear." But in recent years, the perceived relevance of a nuclear test as the definitive marker of a state's nuclear status has been declining. Many scholars and analysts have pointed out that from a technical point of view, an explosive test may not be strictly necessary for the construction of weapons of fearsome and unparalleled destructive power. Therefore, understandably wishing to avoid strategic surprise, they have increasingly shifted their attention to a prior rung on the nuclear proliferation "ladder": a state's accumulation of enough fissile material for a bomb, also known as a significant quantity (SQ). Indeed, no less an authority than Mohamed ElBaradei, former director general of the International Atomic Energy Agency (IAEA), routinely labeled states that accumulated an SQ as "virtual" nuclear weapon states: "countries that are able to develop nuclear weapons overnight."[1] Moreover, there is a strong tendency today to equate "virtual" with "real," given the supposed ease of conversion of fissile materials into weapons before IAEA inspectors can sound the alarm over the new development.[2]

There has never been an authoritative decision to shift the threshold for defining nuclear weapon stateness down from the test/no-test to the SQ/no-SQ indicator. Rather, the shift to SQ/no SQ has been occurring gradually and not even entirely consciously, driven by policy makers trying to deal with concrete diplomatic problems. Not surprisingly, therefore, analysts are inconsistent in their choice of indicator. For instance, Joseph

*Nonproliferation Review, Vol. 17, No. 1, March 2010*
ISSN 1073-6700 print/ISSN 1746-1766 online/10/010161-20

DOI: 10.1080/10736700903484728

Cirincione uses the test/no-test indicator to hail the Treaty on the Non-Proliferation of Nuclear Weapons (NPT) as ''one of the most successful security pacts in history,'' but he also uses the SQ/no-SQ indicator to question the NPT's sustainability when even a treaty-respecting country like Japan can produce enough plutonium to become, at least in some eyes, an ''associate member of the nuclear club.''[3] Although this sort of definitional fuzziness is likely to persist for years, the overall trend over the past two decades has unmistakably been toward the entrenchment of the SQ/no-SQ indicator. Nuclear testing is increasingly seen as a luxury with little real bearing on a state's ability to produce nuclear bombs, at least of the ''simple'' fission variety.

The general trend toward SQ/no SQ can be seen in the evolution of intelligence assessments of the nuclear status of North Korea. For instance, in 1983 the Central Intelligence Agency (CIA) could still write, ''We have no basis for believing that the North Koreans have either the facilities or materials necessary to develop and test nuclear weapons.''[4] By contrast, as detailed by Leon Sigal, ''A National Intelligence Estimate [NIE] in late November 1993 concluded that North Korea already had one or two nuclear weapons,'' despite the fact that no test had taken place.[5] Top Clinton administration officials did attempt to highlight the NIE's precise language of a ''better than even'' chance, rather than a certainty, of actual North Korean nuclear weapons, but this was cold comfort.[6] And after George W. Bush became president, even such faint attempts to maintain some nuance died away. For instance, the CIA under Bush baldly stated in an unclassified 2002 briefing to Congress that the United States ''has assessed since the early 1990s that the North has one and possibly two weapons using plutonium it produced prior to 1992.''[7]

The key point to focus on here is not the progress of North Korea's nuclear program, but rather the changing *metric* used by the intelligence community to assess that progress. In the 1980s, the key question had been whether North Korea would ''develop and test'' nuclear weapons, but by the early 1990s, the intelligence community—having become very uncertain about the test/no-test indicator—was asserting the existence of actual North Korean nuclear weapons simply on the grounds that Pyongyang had allegedly accumulated an SQ with proliferation intent. This shift in metrics was hardly limited to analysts toiling away in the bowels of the CIA headquarters at Langley, Virginia. The notion of an actual North Korean nuclear weapons stockpile was taken so seriously in Washington that prominent liberal and neoconservative analysts agreed that it helped explain why the United States attacked Iraq but not North Korea in the spring of 2003.[8] And even the IAEA's ElBaradei, when pushed in 2004 on whether Pyongyang had actually used its plutonium to produce weapons or merely weapon-grade fuel, replied, ''What's the difference?''[9]

But when North Korea actually did conduct its first test in 2006, it became clear that these experts had been wrong: there was in fact a big difference between having an SQ and having the bomb. North Korea's 2006 test device had a pitiful explosive yield of probably less than 1 kiloton: surely not what Kim Jong Il intended.[10] The failure of the test revealed—unless one gave Pyongyang a huge benefit of the doubt—that as of 2006 North Korea remained a non-nuclear weapon state in spite of itself. Therefore, U.S. policy under

both Bill Clinton and George W. Bush had been based on a gross overestimation of Pyongyang's capabilities. The SQ/no-SQ indicator had seriously failed.

The case of North Korea shows clearly that the question of what indicators we should use to demarcate between nuclear and non-nuclear weapon stateness is one that carries very high analytic and political stakes. This article aims to assess the pluses and minuses of the two most widely used indicators of nuclear weapon stateness—test/no test and SQ/no SQ—from the perspective of *measurement validity*. In other words, the article asks which of the two metrics for evaluating a state's nuclear weapon status is best if the goal is to maximize the accuracy of descriptive inferences. My basic argument is that although the test/no-test indicator clearly has its problems, to replace it with the SQ/no-SQ indicator would be to throw the baby out with the bathwater. I then propose a ''neo-traditional'' approach that maintains the basic test/no-test indicator, but supplements it with a theory-driven analysis of different states' incentives and disincentives to induct nuclear weapons without prior testing.



## Conceptual Background and the Traditional Operationalization

The procedure for measurement validation followed in this article is one that was recently spelled out by the political scientists Robert Adcock and David Collier.[11] In their work, Adcock and Collier highlight the importance of carefully considering the validity of a proposed measurement strategy at three key levels: *conceptualization*, or the formulation of a systematized understanding of the thing to be measured in light of the theoretical meanings and understandings we associate with it; *operationalization*, or the development of specific metrics, or indicators, that clearly reflect the basic concept; and *scoring*, or the application of those indicators to the universe of relevant historical cases. Although this procedure would appear on the surface to be a quite deductive one, in fact Adcock and Collier stress that information from the more ''applied'' steps of the process can (and should) stimulate rethinking at the more ''fundamental'' levels.

According to Adcock and Collier's procedure, it is important first to lay out the conceptual foundations of nuclear weapon stateness. In other words, we must ask: how important is it to draw a clear distinction between ''nuclear weapon states'' and ''non-nuclear weapon states''? Or, for instance, should we instead speak of degrees of nuclear weapons possession, as we speak of degrees of economic advancement? In fact, although perhaps the overall proliferation process can indeed be conceptualized as a ''ladder,'' it has long been recognized that one of the rungs of that ladder must be marked in red. For the acquisition of even a small nuclear weapons arsenal is a revolutionary act in international politics. As such, it engenders uncertain but potentially vast consequences for the acquirer's international position, and indeed for overall international stability.[12] Recognition of the revolutionary implications of the birth of new nuclear weapon states has led the international community to devote enormous diplomatic, financial, human, and other resources to prevent it from happening. Among academics, too, the question of why states ''go nuclear'' receives intense focus, whereas why states ''go aircraft carrier,'' for

instance, does not. In short, conceptually there is an important *qualitative* distinction between the two types of states, non-nuclear and nuclear. Our operationalization of the terms ''nuclear weapon state'' and ''non-nuclear weapon state'' needs to respect this conceptual background.

The traditional focus on the nuclear test as marking the ''birth'' of a new nuclear weapon state fits the above-described underlying concept quite well. A successful nuclear explosive test is typically a clear, focused, and public demonstration of power.[13] The explosive nature of the test closely mirrors our understanding of ''going nuclear'' as a revolutionary act in international politics. This is an important plus for the test/no-test indicator from the perspective of measurement validity.

A second advantage of using the test/no-test indicator is that it greatly reduces the room for the politics and ''spin'' that unfortunately bedevil many measures of key concepts in international relations and even hamper supposedly ''technical'' assessments of the workability of a weapon's design or the quality of its manufacture. It is hard to argue with several kilotons of instantaneously released energy. This is one important reason why the U.S. nuclear weapons laboratories, for instance, have been so reluctant to abandon ''hot'' testing in favor of computer simulations.[14] Of course, small nuclear tests (e.g., North Korea's two efforts to date) do pose problems for interpretation. There is no clear minimum explosive yield that a nuclear test must attain in order to be counted a technical success, but historically a yield of less than about 6 kilotons has tended to give rise to technical doubts. Therefore, for instance, after its 1974 test had a disappointingly small yield—perhaps in the 2–4 kiloton range—Indian nuclear scientists considered a new test prior to induction of weapons as a ''technical necessity.''[15] Still, even a small nuclear test provides a clear signal that a state is at least on the nuclear threshold.

A third advantage of using the test/no-test indicator is quite simply the fact that this is the operationalization of nuclear weapon stateness that has been institutionalized in the foundational text of the nonproliferation regime, the NPT. Article IX of the NPT reads, ''For the purposes of this Treaty, a nuclear-weapon State is one which has manufactured and exploded a nuclear weapon or other nuclear explosive device prior to January 1, 1967.'' To redefine nuclear weapon states as something other than states that have ''manufactured and exploded'' a nuclear device would mean overturning years of settled usage, inevitably leading to dangerous confusion about different states' rights and responsibilities under the treaty. Indeed, given the fundamental position of the non-nuclear versus nuclear weapon state distinction in the construction of the nonproliferation regime, a significant change in our understanding of that distinction implies the regime's complete overhaul. By way of comparison, Article IX also includes the 1967 cutoff date for ''legal'' nuclear weapon state status. Recent attempts to remove India's nuclear pariah status for testing later than 1967—even though it never signed the treaty—have been strongly attacked by nonproliferation advocates as extremely destabilizing to the regime.[16] The nonproliferation community's concerns on this point are quite understandable, but it should therefore be *even more* concerned about the destabilizing consequences of abandoning the NPT's much more central ''manufacture and explode'' criteria for distinguishing between nuclear and non-nuclear weapon states.

The foregoing has mentioned some considerable advantages to the test/no-test metric. However, it must be pointed out that this metric is no more than an *indicator* of a state's potential nuclear weapon stateness, not the underlying quantity of interest itself. In an ideal world, nuclear weapon states would be identified by simply pointing to their actual, operational nuclear weapons, as for instance birds are identified by their feathers. But the combination of the complexity of nuclear weapons systems and state secrecy often makes such a straightforward identification exercise impossible. Pointing to the test is a second-best option, and it should not be forgotten that a nuclear test does not in itself equal a nuclear arsenal. It is actually merely a signal of intent and capacity, not the thing itself. For instance, while India's May 1998 nuclear tests did finally resolve the country's long debate over whether or not to ''go nuclear,'' from an operational standpoint the country still had much work—at least a year's worth—to do before it could credibly claim to own even what Ashley Tellis terms a nuclear ''force-in-being'': ''a deterrent whose effectiveness derives from its ability to be constituted into a viable retaliatory instrument under conditions of supreme emergency.''[17] In other words, if we liken nuclear weapons to a fountain pen, India's 1998 tests proved that it had an adequate ink reservoir, but it still needed some time to add the nib that is essential for writing.

In short, using the test as the measure of nuclear weapon stateness actually jumps the gun in most circumstances. But the evident imperfection of this indicator should not immediately lead us to condemn it as invalid for making descriptive inferences. All operationalizations of key concepts in human affairs are imperfect. The proper standard for indicators in the social sciences is not perfection, but adequacy. Furthermore, on a more practical level, given the destabilizing consequences of strategic surprise, it can be argued that it makes sense to choose an indicator that alerts the members of the international community to the reality of a new nuclear weapon state slightly in advance, while they still have some time for cool deliberation on how best to adjust their strategic postures in light of the emergent reality.

A successful nuclear test clearly indicates nuclear weapon stateness because it mirrors our underlying concept, sends a relatively unambiguous signal of the state's nuclear attainment, and has for a long time been part of diplomatic praxis. But is the *absence* of a nuclear test a valid indicator of *non*-nuclear weapon stateness? If it is not, then as much as we might like to maintain the test/no-test indicator for the above-mentioned reasons, in fact it is an invalid metric and needs to be replaced.

In fact as many have pointed out, there is an evident problem with using the no-test criterion as the indicator of non-nuclear weapon stateness. The problem is that from a purely technical point of view a nuclear weapons test is not always strictly necessary for nuclear bombs to be built. For example, the ''Little Boy'' gun-type uranium bomb was never explosively tested before the United States dropped it on Hiroshima. Even so, during the first decades of the nuclear age, few doubted that the test was the inevitable culmination of a nascent nuclear weapons program. All of the first five nuclear weapon states carried out explosive tests before inducting nuclear weapons into their military arsenals. Even in the U.S. case, its first nuclear blast was not over Hiroshima, but instead near Alamogordo, New Mexico, some three weeks earlier. And despite the different design of the Alamogordo and Hiroshima devices (the former being of a ''Fat Man''–type

plutonium implosion bomb), without the Alamogordo test U.S. scientists and policy makers would have been unconvinced that the bomb was indeed in their hands and that they could therefore order its use against Japan.[18] The consensus assumption that the birth of new nuclear weapons arsenals required prior nuclear testing was what led NPT Article IX to be written as it was.

This consensus emerged not only because of the behavior of the first five nuclear weapon states, but also because that behavior conformed to the tenets of deterrence theory as it had developed by the mid-1960s.[19] As was memorably explained to the Soviet ambassador in the 1964 movie *Dr. Strangelove*, deterrence theory's basic insight was that in order for the apocalypse to be avoided, its potential arrival must be heavily advertised.[20] The nuclear test is perfectly suited for that task. Indeed, it is the only sure means of lending an air of reality to the otherwise "fabulously textual" world of nuclear strategy.[21] In short, from the perspective of mainstream deterrence theory, if you don't test, then you don't have the bomb—even if you do have it. So therefore you must test.

During the Cold War, when it came to defining nuclear weapon states this *strategic* argument that testing was strictly necessary was generally seen to trump the *technical* argument that it was not strictly necessary. But today, many are making a different calculation.



## Critiques of the Test/No-Test Indicator

In recent years, the utility of using the first nuclear test as the key indicator of a state's nuclear weapon status has come under increasing attack. Although critics of the traditional thinking accept that a nuclear test reflects a state's entry into the weapons state club, they argue that the lack of a test does not mean that the state has *not* joined that club. In other words, in the phrase made famous by American neoconservatives, absence of evidence is not evidence of absence. Therefore, there is a growing belief that the real red line between nuclear and non-nuclear weapon states should actually be drawn at the acquisition of an SQ of fissile material.

Why are new-school nonproliferation advocates so skeptical about the test/no-test indicator's ability to capture the reality of proliferation? First of all, their point of view reflects the remarkable post-Cold War decline in the influence of mainstream deterrence theory.[22] As noted above, it was always understood that from a purely technical standpoint, testing was not strictly necessary in all cases; however, this potential problem was thought to be overridden by the strategic necessity of testing. New-school nonproliferation advocates, by contrast, are not at all convinced by the traditional notion that since deterrence theory requires testing, testing will happen. Instead they pose a question: wouldn't you rather be safe than sorry? Behind this deceptively simple question there stands an entire alternative strategic theory, which was originally elaborated by the neoconservatives' godfather, Albert Wohlstetter. On the definition of nuclear weapon stateness, Wohlstetter (and his coauthors) had this to say:

> If, in fact, technological transfers can bring a ''nonnuclear weapon state'' within weeks, days or even hours of the ability to use a nuclear explosive, in the operational sense that ''nonnuclear weapon state'' will have nuclear weapons. The point is even more fundamental than the fact that effective safeguards mean timely warning. A necessary condition for timely warning is that there be a substantial elapsed time. But if there is no substantial elapsed time before a government may use nuclear weapons, in effect it *has* them.[23]

Wohlstetter et al. continued: ''Consider the case of a government which is not at war, but is capable of quickly assembling a nuclear device to use or threaten to use against another government without such a capability. Once again, there is no practical difference between the coercion it could use or the threat it could execute from what a nuclear power might manage.''[24]

In short, whereas the traditional deterrence theory view was that if you might not have the bomb then you don't have it, the Wohlstetter view was that if you might have it then you do have it. Thus, he saw the necessity of shifting the indicator of nuclear weapon status from test/no test to SQ/no SQ.

Wohlstetter's perspective has been updated and applied to the question of the reform of the NPT regime in a recent volume edited by the well-known nonproliferation advocate Henry Sokolski. Sokolski and his colleagues explicitly call for fundamentally reinterpreting the NPT not merely as a ban on non-nuclear weapon state acquisition of a nuclear weapons arsenal, but in fact as a ban on their acquisition of an SQ, or indeed even of facilities (such as uranium enrichment or fuel reprocessing plants) that can readily produce an SQ in a short period of time.[25] The argument runs as follows: safeguards or no safeguards, timely warning of conversion of an SQ into operational weapons is impossible; therefore, outsiders cannot help but treat a state that has an SQ as if it already had an operational weapons arsenal. Furthermore, because under the NPT non-nuclear weapon states have taken on the obligation not to acquire nuclear weapons, and since in terms of its strategic implications an SQ is actually equivalent to a nuclear weapon, therefore non-nuclear weapon states are also obligated by the treaty not to acquire an SQ. In short, the decision to seek to acquire an SQ is tantamount to a decision to ''go nuclear.''

New-school nonproliferation advocates also base their case on an empirical perception that some states—and perhaps even many—actually prefer to forego testing before inducting nuclear weapons into their arsenals. For instance, Israel has been armed with operational nuclear weapons probably since 1973, and perhaps as early as 1967.[26] Indeed, most expert observers believe that Israel today has not merely ''the bomb,'' but indeed a large and technically sophisticated nuclear arsenal mounted on multiple delivery systems. Yet in contrast to all of the other nuclear weapon states, Israel has apparently never carried out an explosive nuclear test.[27] Moreover, it refuses even to officially acknowledge the existence of its arsenal, although there have been many ostensible ''slips of the tongue'' over the years.[28] The Israeli case therefore appears to represent a major exception to the ideal-typical nuclear path that was followed by the first five nuclear weapon states. Note that the problem posed by the Israeli case for traditional

understandings of the natural course of proliferation is not that it developed the bomb in secret, but rather that having built the bomb, it *maintained* the secret.

The mere fact of the Israeli exception is not in itself damning to the test/no-test indicator. After all, no indicator is 100 percent error-free. If one had good reason to believe that Israel's no-test path to the bomb was *sui generis*, one of a kind and unlikely to be repeated, then the Israeli exception would be just that, an exception to the rule. And indeed, for many years proliferation experts quite happily retained the test/no-test indicator in general, while simply appending an asterisk to the Israeli case.

However, if the measurement error introduced by an indicator is not merely minor and random but large and systematic—in other words, if the indicator is seriously *biased*—then the basic validity of the indicator does come into question.[29] And especially since the early 1990s, scholars and policy makers have developed the argument that the Israeli case is indeed not *sui generis,* but is rather at the vanguard of a worldwide trend away from testing as the key step across the threshold to nuclear weapons acquisition. In a seminal work Avner Cohen and Benjamin Frankel characterized the Israeli case of "opaque proliferation" as *ideal-typical* of a burgeoning second generation of new nuclear weapon states, which was being sorely underestimated by the traditional test/no-test indicator.[30] Cohen and Frankel also pointed to Pakistan and South Africa as states that had built "opaque" nuclear arsenals.

Over the past two decades the Cohen-Frankel opaque proliferation model has become increasingly accepted in the policy world (though less so among academics) as a true picture of contemporary reality. Thus we are often informed that we are now living in a "second nuclear age" of rampant, more-or-less hidden nuclear proliferation.[31] And the response by policy makers to this perceived problem, as noted in the introduction to this article, has been to implicitly move back the red line from the first test to the production of an SQ of fissile material—the last chance, they believe, to catch the proliferant state red-handed before it becomes capable of a strategic surprise. The next section evaluates this SQ/no-SQ alternative to the traditional test/no-test indicator of nuclear weapon stateness.



## Critiques of the SQ/No-SQ Indicator

As noted above, the emerging operationalization of the concept of nuclear weapon stateness focuses on the acquisition of an SQ of fissile material, which some see as a more valid indicator of de facto nuclear weapon state status than the traditional test/no-test indicator. Does this argument compute? Should we choose definitively to abandon the test/no-test indicator in favor of the SQ/no-SQ alternative? In fact, there are several serious drawbacks to this idea. First, it is questionable to assert that the Israeli posture of nuclear opacity is appealing to the new generation of nuclear weapon states. Second, because the accumulation of an SQ is much less visible than a nuclear test, using that metric to assess nuclear weapon status introduces much greater uncertainty and even opens the door to politically biased misinterpretation. Third, the SQ/no-SQ indicator considerably underestimates the obstacles, both technical and political, that stand between the accumulation

of fissile material and the production of operational bombs. And last but not least, shifting to the SQ/no-SQ indicator may have a significant impact on real-world behavior, and quite possibly could have the effect of encouraging more nuclear proliferation. The article will flesh out these points one at a time.

## *Israel as a* Sui Generis *Proliferator*

First, as noted above, Israel has been showcased as the ideal-typical case of "opaque proliferation," but is the Israeli path really ideal-typical of a larger class of cases, or is it just an anomalous exception? One way to begin answering this question is to ask whether Israel's stance of nuclear opacity has proven to be a smart choice that others might therefore be tempted to follow. It is certainly possible to argue that the Israeli bomb has proven key to the country's security.[32] But Israel's consignment to the realm of nuclear opacity has also been a difficult experience in many ways. Opacity has led to unclear nuclear signaling externally, a stunted nuclear doctrinal debate internally, regular embarrassment at international diplomatic forums, and even the cancerous growth of a secret state that undermines Israeli democracy.[33] Indeed, Israel's leaders often appear, at least at some level, to want to put an end to the ambiguity—as evidenced by their numerous ostensibly inadvertent admissions about the fact of the Israeli bomb over the years.[34]

It is hard to imagine that many other states would be eager to join Israel in this dubious club.[35] And indeed most of the other states that were once viewed as part of it—notably India, North Korea, Pakistan, and South Africa—clearly are not in it today.[36] Even Avner Cohen, the coauthor of the seminal work on "opaque proliferation," now argues that Israel is a "*sui generis* proliferator," whose choice to rely for decades on an untested nuclear arsenal is historically unique.[37]

## *SQ or No SQ? The Difficulty of Assessment*

A second problem with the SQ/no-SQ indicator lies in the question of how we can know that an SQ has actually been achieved. For instance, as was noted at the outset of this article, the U.S. intelligence community believed that North Korea had at least one plutonium-based bomb by the early 1990s. Moreover, in 2002 the United States publicly alleged that Pyongyang had also developed a uranium enrichment capability with the help of the Pakistani scientist A.Q. Khan. But it now appears that the country actually did not have a secret SQ of plutonium, let alone nuclear weapons, in the 1990s. As part of the six-party talks process, in May 2008 North Korea turned over around 18,000 pages of technical documents to the United States and other negotiating partners. According to reports, these documents—which would be very difficult to fabricate—reveal that contrary to U.S. estimates, North Korea had removed no plutonium from its Yongbyon reactor prior to 1992. In other words, the experts now believe that the assumption that North Korea had an SQ, let alone actual nuclear bombs, already in the early 1990s was

simply wrong.[38] Moreover, the equally widespread notion that, thanks to A.Q. Khan, North Korea may have been on the brink of acquiring a considerable stockpile of HEU at the start of the 2000s also now appears quite suspect. The U.S. intelligence community's uranium enrichment assessment at that time turns out to have been based largely on North Korea's import of the same type of aluminum tubes that led to wildly exaggerated estimates of the progress of Iraq's nuclear program in the early 2000s.[39] Of late, the United States has been remarkably quiet on the North Korean uranium enrichment issue.

The retreat of the analysts from their prior strong statements about the advancement of the North Korean nuclear program is not just a proper correction in light of new evidence. Rather, it shows a deeper flaw in the reliance on the SQ/no-SQ indicator as the measure of nuclear weapon stateness. The validity of any indicator depends heavily on the degree of measurement error that it introduces. Yet the application of the SQ/no-SQ indicator requires a great deal of guesswork, even if the state's nuclear facilities are known and easily observable via satellite. If there are suspicions of hidden facilities, the ratio of guesswork to known facts becomes much more lopsided—and this ratio increases yet further if we entertain the possibility that a state might have acquired fissile material from Russian or other smuggling rings. Therefore, the SQ/no-SQ indicator seems to open the door widely to all manner of interpretations of the nuclear progress of particular cases, and therefore also to open it to political abuse of the sort that characterized the reports on Iraq's alleged WMD programs in the run-up to the 2003 Iraq war.[40]



## *From an SQ to the Bomb: A Difficult Road*

An additional serious drawback to the SQ/no-SQ indicator is the assumption of easy conversion to bombs on which it rests: as the U.S. official Laura Holgate put it, "Once you have the fissile material, it's a matter of basic chemistry, basic machinery, and a truck."[41] This grossly underestimates the degree of technical difficulty of putting all the pieces of the puzzle together, and perhaps even more importantly it also ignores the political and organizational difficulties that can hamper and delay the shift from bomb program to weapons arsenal.[42] For these reasons, even if we are sure that an SQ has been achieved and that the country's leadership intends to go all the way, the SQ/no-SQ indicator is still prone to overstate the number of bona fide nuclear weapon states.

The mistake of assuming that the road from SQ to bomb is a short and easy one is again highlighted by the misbegotten U.S. experience with North Korea. Former Assistant Secretary of State for East Asian and Pacific Affairs James Kelly, who had earlier happily buried the U.S.–DPRK Agreed Framework, told the press shortly after leaving office in 2005 that since it was clear that North Korea already had the bomb, "I personally have never been worried about a test. A test would just tell us what we already know."[43] Yet as noted above, in fact when the first North Korean nuclear test actually happened in October 2006 it was quite underwhelming. The blast measured somewhere between 0.2 and 1 kiloton in yield—making it by far the least powerful first nuclear test in history. Thus, far from confirming the DPRK as a nuclear weapon state, the test actually suggested that the

opposite was true. As Jungmin Kang and Peter Hayes of the Nautilus Institute put it after reviewing the technical evidence, "We conclude that the DPRK test did not enable it to pole-vault into the ranks of nuclear weapon states. . . . The DPRK inhabits a peculiar and ambiguous status between having declared itself to be nuclear-armed, and having demonstrated by its test that it is not capable of such armament."[44] North Korea's claimed second nuclear test, in 2009, appears to have been somewhat more successful, but doubts about it persist as well.[45] Moreover, even if the second test met the minimum standards for success, there would still be an open question about North Korea's capacity to take the next step and produce reliable, deliverable nuclear warheads.

### *Real-World Effects of the Shift to SQ/No SQ*

Finally, we must not merely consider the validity of shifting to the SQ/no-SQ indicator on conceptual and empirical grounds, but also what consequences such a shift might have in the real world. After all, unlike the natural sciences, in the social sciences we must be mindful that our theories, concepts, and measurement strategies may not merely describe but also influence the behavior of the subjects under study.[46]

Indeed, one of the main motivations for those who are attempting to install the SQ/no-SQ indicator is precisely the hope that the widespread use of such a metric would change state behavior. They reason that if states can be convinced that the choice to make fissile material actually itself represents the choice to "go nuclear," then many fewer states will in fact end up going even that far down the nuclear path; meanwhile, others' true intent will be unambiguously identified sooner.[47] That would be a boon for international security. But although this is a plausible story, other stories are plausible, too. For instance, the shift to SQ/no SQ represents a significant "dumbing down" of the barriers to entry into the nuclear "club." Why should this not encourage more states to try to get in? Indeed, why should it not lead even frankly incompetent states like Zimbabwe or Myanmar to imagine that with just a little nuclear contraband they, too, would be able to claim nuclear weapon state status? Moreover, might not states that already have an SQ, such as Japan, simply accept the world's new judgment of their nuclear weapon stateness, despite never having made that decision for themselves? After all, a key factor encouraging nonproliferation up to the present day has been state leaders' perception that to "go nuclear" is a revolutionary choice, with potentially vast and unknowable consequences for their states' security and power positions.[48] But if the international community pushes the definition of nuclear weapon stateness backward to SQ/no SQ, it will essentially have relieved some states' leaders from the burden of facing that revolutionary choice.

Furthermore, even the mere *effort* to replace test/no test with SQ/no SQ could be harmful to overall international systemic stability. For although the lines we draw between two qualitatively different states of being may be social constructions, it does not logically follow that these lines can be drawn arbitrarily or unilaterally.[49] The test/no-test indicator became the red line between non-nuclear and nuclear weapon state status because, in light of our background understandings of the nuclear issue, it was the *obvious* focal point—and it still is.[50] In politics, and particularly in international politics where the

dangers of miscommunication are rampant, the wise course is usually to align our concepts and indicators as closely as possible in accordance with ordinary usage. Attempts to change by fiat the common web of understandings at the very foundations of international order are likely to produce confusion, or even chaos.[51] In short, the SQ/no-SQ indicator for nuclear weapon stateness is not only analytically flawed, but it is also politically dangerous.

None of the foregoing should be taken to imply that efforts to enhance the tracking or to control the growth of fissile material stockpiles are meaningless or self-defeating. Rather, such efforts may well be an important means of promoting nonproliferation.[52] But still, fissile material is not ''the bomb,'' and to equate the two is a fundamental error in terms of both social science and public policy.



## From *Whether* Test or No Test to *Why* Test or No Test?

The previous section demonstrated the many problems from the perspective of measurement validity of moving wholesale from test/no test to SQ/no SQ as the preferred indicator of nuclear weapon stateness. In light of these problems, it makes much more sense to retain the traditional test/no-test indicator, which has actually not served us all that badly up to now. However, the fact that the simple test/no-test indicator is superior to the usual proposed alternative does not mean that we should be completely happy with it. As previously discussed, new-school nonproliferation advocates are certainly right to note that the simple test/no test-indicator is not infallible, and they could possibly be right that the indicator will prove more fallible in the future.

So, can we do better? In their important work on measurement validity, Adcock and Collier propose a method of ''nomological'' or ''construct'' validation, which they also catchily term the ''AHEM'' method of validation—Assume the Hypothesis, Evaluate the Measure.[53] In other words, Adcock and Collier suggest that when we are confronted with the task of measuring something that we cannot see directly, one possible response is to look for the presence or absence of the established *causes* of that thing. Where those causes are seen to be present, we can make at least a cautious claim that their typical effect should also be present.

An AHEM approach to measurement validation in this context leads us to ask not just *whether* a state has conducted a test, but also *why* states test, or not, at the outset of their careers as nuclear weapon states. In other words, although the test/no-test indicator has historically done pretty well as a measure of nuclear weapon stateness, the truth is that we do not know why. As noted previously, the traditional confidence in the indicator was based on the assumption that states should instinctively understand and comply with the requirements of deterrence theory. But in light of U.S. and Soviet nuclear behavior during the Cold War, it is hard to claim that states really act as traditional deterrence theory imagines they do.[54] The test/no-test indicator has thus actually fared better empirically than it should have, given its heavy debt to this rather dubious theory of state

behavior. This lack of proper theoretical grounding was what opened the door for advocates of the SQ/no-SQ alternative in the first place.

A more theoretically grounded, AHEM approach to operationalizing the concept of nuclear weapon stateness would also improve on the current test/no-test indicator by allowing us to report our level of confidence in the validity of that indicator, both globally and on a case-by-case basis. After all, valid measurement means not merely making careful descriptive inferences, but also carefully estimating the degree of uncertainty attached to those inferences.[55] This is always a difficult task, but it is much more difficult if we have not first identified the main factors that drive countries to choose the route of testing or not testing. Paradoxically, by delineating the key potential problem cases for the test/no-test indicator, the AHEM approach should actually strengthen our faith in the overall value of that indicator.

A full elaboration of the AHEM approach would require a book-length effort. But to provide a flavor of the approach, I propose the following hypothesis as a possible starting point: *Civilian-run nuclear weapons programs are highly unlikely to induct nuclear weapons without a test, but military-run nuclear weapons programs may sometimes choose to do so.*

To understand the thinking behind this hypothesis, first note that the decision to induct nuclear weapons without prior testing most likely reflects a highly offensive strategic orientation. This is because by not testing, the state is maximizing its potential for strategic surprise. Strategic surprise can be had even if the nuclear weapons program is not a complete secret, for without a test the adversary will be unaware of the true advancement of the arsenal, and it may even be lulled into complacency by official denials. In short, induction without prior testing is not a sign of restraint, but rather quite the opposite. The state may be readying its unannounced nuclear arsenal for a ''splendid first strike,'' or it may have other plans, such as to use a stunning test of operational, battlefield nuclear weapons to put the enemy on warning, as South Africa was ready to do during the 1980s.[56] But even the latter would clearly be an extremely aggressive maneuver.

If inducting without testing reflects a highly offensive orientation, then who in the state is likely to support such a policy? It is widely claimed in the international relations literature that professional military organizations, in contrast to most civilian politicians and agencies, have a decided bias in favor of offensive strategies. Various reasons have been adduced for this preference: military organizations perform better when they can follow their own script; militaries embrace offensive strategies as a means of gaining larger budgets and operational autonomy; military professionals tend to be fixated on, and therefore to overestimate, the likelihood of war; military culture tends to glorify offensive daring; militaries are accustomed to maintaining the secrecy required for launching first strikes; and so forth.[57] In short, many scholars concur that the military preference for the offensive is a strong, enduring, global tendency. (Beyond professional militaries, the most likely candidates to display similar tendencies would probably be paramilitary or secret police organizations.) As evidence for this tendency in the nuclear area, despite numerous claims by analysts that nuclear weapons are the defensive weapon *par excellence,* various states' militaries have tried hard to employ nuclear weapons in strategies of compellence as well as deterrence, and many have also shown a clear inclination toward nuclear war-fighting doctrines.[58]

Since militaries tend to prefer offensive strategies, and since induction without testing is a highly offensive strategy, the natural conclusion is that militaries should be more likely than civilians to do without testing. However, ''more likely'' here should not be taken to mean ''likely.'' Even offensive-minded militaries may still prefer first to garner the presumed deterrence benefits of the bomb by carrying out a successful test, and only then turn their minds to thinking about possible offensive uses of the weapon.[59]

Moreover, in addition to their interest in deterrence, militaries also have a strong desire to know that the weapons in their arsenal will actually work. This need is especially acute if nuclear weapons are indeed slated to be employed offensively. After all, if the plan is to carry out a splendid first strike, that first strike had better be truly splendid. This desire for reliability of the nuclear arsenal also militates in favor of testing.

One way for the military to satisfy its desire for reliability while maintaining the option of strategic surprise could be to carry out a secret test. This was the option pursued by the United States and apparently also by the Soviet Union, although the Soviet blast was quickly discovered by outside intelligence.[60] The option of carrying out a secret test remained possible as late as 1979, when South Africa may have detonated a device in the South Atlantic.[61] But due to the dramatic increase since then in quantity and quality of remote-sensing equipment of various kinds (especially the largely deployed International Monitoring System for verifying compliance with the Comprehensive Nuclear-Test-Ban Treaty), a secret test is clearly not an option anymore. If the United States and international community could ''hear'' even North Korea's nuclear fizzle of 2006, then the option of a secret test is clearly no longer on the table. So if the military wants to go for nuclear opacity today, it has to do without a real test. This fact considerably raises the costs of pursuing such an offensive nuclear strategy.

I have argued that unlike most civilian politicians and agencies, militaries may be tempted to induct nuclear weapons without prior testing. The question now becomes, under what circumstances are those military preferences more likely to be realized? This will of course depend largely upon the specific institutional balance of power inside the state. In general it may be surmised that if the military controls the nuclear weapons program—i.e., either directly runs the program, or is directly and uniquely reported to by those who do run the program—it is likely to have a disproportionate say over the question of whether or not to test prior to induction. This is because it would be in a position to monopolize the information flow to the other parts of the state—and thus to twist the information toward or away from the testing option as it saw fit. We thus arrive at the hypothesis stated above, that while civilian-run nuclear weapons programs are highly unlikely to induct nuclear weapons without a test, military-run nuclear weapons programs may sometimes choose to do so.

A brief study of the historical record appears to provide tentative evidence in favor of this hypothesis. Before turning to that historical record, however, it must be emphasized once again that most nuclear weapons programs have never gotten to the stage where they were faced with the decision to test or to induct without testing. For instance, North Korea's failure to test nuclear weapons in the early 1990s was not a sign of a strategic choice for nuclear opacity, but rather simply a reflection of the program's lack of sufficient technical advancement.

Since 1945, ten states are generally believed to have crossed the line and either tested a nuclear fission device, or inducted untested but operational nuclear weapons into their arsenals. Table 1 below shows that the hypothesized distinction between

**TABLE 1**

Nuclear testing choices of military- versus civilian-run nuclear weapons programs.

| | **Military-run** | **Civilian-run** |
|---|---|---|
| **Test, then induction** | China, North Korea | United Kingdom, France, India |
| **Secret test, then induction** | United States, possibly South Africa | Soviet Union |
| **No test prior to induction** | Israel, Pakistan | n/a |

**Key:**

*United States:* Program run by the military with civilian input. First test in July 1945, prior to induction, although the Hiroshima attack employed a different, untested bomb design. Test carried out in total secrecy and results not announced until after Hiroshima.[1]

*Soviet Union:* Program run by state security apparatus (the KGB). First test in 1949, prior to induction. Test secret but aboveground and quickly detected by U.S. intelligence.[2]

*United Kingdom:* Program run by civilian atomic energy authority with military input. First test in 1952, prior to induction. Test results announced.[3]

*France:* Program run by civilian atomic energy authority with military input. First test in 1960, prior to induction. Test results announced.[4]

*China:* Program run by the military. First test in 1964, prior to induction. Test results announced.[5]

*Israel:* Program run by the General Staff of the Defense Forces, reporting to civilian leadership in the Ministry of Defense. Weapons inducted without testing, although possibly participated in the supposed South African test.[6]

*India:* Program run by civilian atomic energy authority with some military input. First test in 1974, prior to induction. Test results announced. Second test in 1998 also probably prior to induction, although the country may have had a crude operational capacity by then. These test results also announced.[7]

*South Africa:* Program run by civilian atomic energy authority until 1979, when military parastatal Armaments Corporation of South Africa took over. Possible test in 1979, prior to induction. No announcement. Secrecy maintained until weapons dismantled in the early 1990s.[8]

*Pakistan:* Program run by the military after 1976. According to U.S. intelligence, probable induction in 1990, prior to testing. First test in 1998, results announced.[9]

*North Korea:* Program run by civilians reporting directly to the top political leadership at least until early 1990s.[10] After that point great uncertainty, but given the regime's overall "military-first" policy the military has likely taken over the project, particularly after 2002. First test in 2006 had poor results, though regime publicly announced it as a success. Second apparent test in 2009, with seemingly slightly better results, again announced. Induction almost certainly not before 2006, and indeed serious questions about operational capabilities remain unresolved.

[1] See Richard Rhodes, *The Making of the Atomic Bomb* (New York: Simon and Schuster, 1986).
[2] See David Holloway, *Stalin and the Bomb: The Soviet Union and Atomic Energy, 1939–1956* (New Haven: Yale University Press, 1996); Jeffrey Richelson, *Spying on the Bomb: American Nuclear Intelligence from Nazi Germany to Iraq and North Korea* (New York: W. W. Norton, *2006),* ch. 2.
[3] See Margaret Gowing with Lorna Arnold, *Independence and Deterrence: Britain and Atomic Energy, 1945–1952* (New York: St. Martin's Press, 1974).
[4] See Marcel Duval and Dominique Mongin, *Histoire des forces nucléaires françaises depuis 1945 [History of French Nuclear Forces since 1945]* (Paris: Presses Universitaires de France, 1993).
[5] See John Lewis and Xue Litai, *China Builds the Bomb* (Stanford, CA: Stanford University Press, 1991).
[6] See Avner Cohen, *Israel and the Bomb* (New York: Columbia University Press, 1998); Helen E. Purkitt and Stephen F. Burgess, *South Africa's Weapons of Mass Destruction* (Bloomington, IN: Indiana University Press, 2005).
[7] See George Perkovich, *India's Nuclear Bomb: The Impact on Global Proliferation,* updated ed. (Berkeley, CA: University of California Press, 2002); Ashley Tellis, *India's Emerging Nuclear Posture: Between Recessed Deterrent and Ready Arsenal* (Santa Monica, CA: RAND, 2001).
[8] See Purkitt and Burgess, *South Africa's Weapons of Mass Destruction;* Peter Liberman, "The Rise and Fall of the South African Bomb," *International Security* 26 (Autumn 2001), pp. 45–86.
[9] See Feroz Hassan Khan and Peter R. Lavoy, "Pakistan: The Dilemma of Nuclear Deterrence," in Muthiah Alagappa, ed., *The Long Shadow: Nuclear Weapons and Security in 21st Century Asia* (Stanford, CA: Stanford University Press, 2008), pp. 215–40; Richelson, *Spying on the Bomb,* esp. pp. 330–32, 345–46.
[10] See James Moltz and Alexandre Mansourov, *The North Korean Nuclear Program: Security, Strategy, and New Perspectives from Russia* (New York: Routledge, 2000); Jacques E. C. Hymans, "Assessing North Korean Nuclear Intentions and Capacities: A New Approach," *Journal of East Asian Studies* 8 (May-August 2008), pp. 259–92.

civilian- and military-run programs does seem to matter in determining how those ten states decided to cross the line: with a test, secret test, or no test.

As Table 1 clearly shows, most states have chosen to test in advance of acquiring operational nuclear weapons arsenals. Moreover, all programs run by civilian agencies have chosen to test before inducting nuclear weapons, though notably the Soviet nuclear program, run by the KGB secret police organization, chose not to immediately announce the test's success. Military-run programs have shown a somewhat greater willingness to do without testing, but even in that category at least three out of six tested prior to induction; four out of six if South Africa indeed conducted a test in September 1979.[62] (Note that although South Africa's program was initially civilian-run, by the time of the test in September 1979 control of the program had passed to the military parastatal Armaments Corporation of South Africa.[63]) A more thorough empirical study might suggest different codings for this table—e.g., the Soviet Union may have assumed the U.S. would discover and announce its test for it; Israel may have participated to some degree in South Africa's supposed test; or North Korea's program may still be civilian-run—but such coding changes would not substantially change the basic conclusion.

What the table does not show is that of the two states that apparently did not test prior to induction, Israel and Pakistan, only Israel definitely committed itself to developing a sizable arsenal without testing. By contrast, whatever genuine operational capability, if any, Islamabad had during its brief period of nuclear opacity was almost certainly very limited.[64] Again, the Israeli case seems to be *sui generis*. It might also be noted that in both the Israeli and Pakistani cases, U.S. intelligence learned about their induction of nuclear weapons very shortly after it had occurred.[65] Therefore the United States, at least, was not vulnerable to strategic surprise in either case.

The tentative takeaway from this empirical exercise is that it makes sense to retain the test/no-test indicator as our basic metric for nuclear weapon stateness, unless there is direct and incontrovertible evidence that successful induction without testing has taken place in a given case. But it also makes sense to append an asterisk to the non-nuclear weapon state status ascribed to military-run nuclear weapons programs that have built up an SQ of fissile material. ''Appending an asterisk'' means retaining the test/no-test indicator but indicating *less than complete confidence* in this judgment. The precise level of confidence for a given case could then be fine-tuned through in-depth qualitative analysis of the strategic culture of the military running the program, and in particular of the degree to which it has generally been prone to exploit opportunities for strategic surprise. It would also be important to study the overall quality of management of the nuclear weapons program, in order to assess the state's ability to move from an SQ to an operational arsenal that, despite remaining untested, could be counted upon to function as planned. On this point it bears repeating that the technical bar for inducting operational nuclear weapons is much higher than the bar for testing a simple nuclear fission device.

In short, the simple test/no-test indicator is always preferable to the SQ/no-SQ indicator as a first-cut approximation of the state of play, even for military-run nuclear weapons programs; but a *better-theorized* test/no-test indicator is much more preferable.

The key lesson here is less the specific hypothesis of civilian- versus military-run programs, and more the overall approach to this knotty problem of measurement validity.

The AHEM, or what one might term ''neo-traditional,'' approach to measuring nuclear weapon stateness provides us with a way forward beyond the test/no-test versus SQ/no-SQ impasse. The approach is ''traditional'' because it retains the focus on the testing stage as the key threshold point. But the approach is also ''neo'' because it admits that the traditional barefoot empiricist indicator of test/no-test is not enough, and indeed that it may even contain a systematic bias toward false negatives in a subset of cases.

The neo-traditional approach to the measurement of nuclear weapon stateness will not be sufficient in itself to produce a consensus coding of the nuclear weapon status of a country such as Iran. But it can help to set the parameters for a more analytically sound debate on Iran or other states' nuclear evolution, and also—even especially—it should reassure us about the continuing overall health of the nonproliferation norm.



## ACKNOWLEDGEMENTS

The author thanks Rieko Kage, John Mueller, Gaukhar Mukhatzhanova, Benoît Pélopidas, William Potter, Maria Rost Rublee, Scott Sagan, and participants at the ''Forecasting Nuclear Proliferation'' project workshop in summer 2008 for their helpful comments on an earlier version of this article. A slightly different version of this article will appear in William Potter and Gaukhar Mukhatzhanova, eds., *Forecasting Nuclear Proliferation* (forthcoming 2010, Stanford University Press).

## NOTES

1. Caterina Dutto, ''ElBaradei Remarks at Georgetown University,'' Proliferation Analysis, Carnegie Endowment for International Peace, October 24, 2006, <www.carnegieendowment.org/publications/index.cfm?fa=view&id=18816>. My critique of ElBaradei's focus on the ''virtual'' NWS status of states with an SQ should not be confused with alarmist critiques of the careful distinction ElBaradei makes between ''virtual'' and ''latent'' NWS. On the latter point see Scott D. Sagan, ''Nuclear Latency and Nuclear Proliferation,'' in William Potter and Gaukhar Mukhatzhanova, eds, *Forecasting Proliferation* (Stanford, CA: Stanford University Press, 2010).
2. Thomas B. Cochran, ''Adequacy of IAEA's Safeguards for Achieving Timely Detection,'' in Henry D. Sokolski, ed., *Falling Behind: International Scrutiny of the Peaceful Atom* (Carlisle, PA: U.S. Army War College Strategic Studies Institute, 2008), pp. 121–58.
3. Joseph Cirincione, *Bomb Scare: The Past and Future of Nuclear Weapons* (New York: Columbia University Press, 2007), pp. 128, 105.
4. CIA, Directorate of Intelligence, ''A 10-Year Projection of Possible Events of Nuclear Proliferation Concern,'' marked ''Secret,'' May 1983, p. 5. Document 2 of Robert A. Wampler, ed., *North Korea and Nuclear Weapons: The Declassified Record,* National Security Archive Electronic Briefing Book No. 87, April 25, 2003, <www.gwu.edu/~nsarchiv/NSAEBB/NSAEBB87/>.
5. Only the State Department's Bureau of Intelligence and Research dissented from the NIE's conclusion. Leon Sigal, *Disarming Strangers: Nuclear Diplomacy with North Korea* (Princeton, NJ: Princeton University Press, 1998), p. 90.
6. Ibid., pp. 91–92.
7. CIA assessment provided to Congress, November 2002. Document 22 of Wampler, *North Korea and Nuclear Weapons*.
8. See, e.g., Fred Kaplan, ''The Unspeakable Truth,'' Slate.com, January 7, 2003, <www.slate.com/id/2076468/>; Karen Elliott House, ''The Lesson of North Korea,'' *Wall Street Journal,* January 3, 2003, p. A10.

9. Quoted in David E. Sanger, "When a Virtual Bomb May Be Better than the Real Thing," *New York Times*, December 5, 2004, <www.nytimes.com/2004/12/05/weekinreview/05sang.html>.
10. Jacques E.C. Hymans, "Assessing North Korean Nuclear Intentions and Capacities: A New Approach," *Journal of East Asian Studies* 8 (May–August 2008), pp. 259–92; Jacques E.C. Hymans, "Discarding Tired Assumptions About North Korea," *Bulletin of the Atomic Scientists* online edition, May 28, 2009, <www.thebulletin.org/web-edition/op-eds/discarding-tired-assumptions-about-north-korea>.
11. Robert Adcock and David Collier, "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research," *American Political Science Review* 95 (September 2001), pp. 529–46.
12. Robert Jervis, *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon* (Ithaca: Cornell University Press, 1989).
13. To assess whether there has or has not been a test is actually very difficult and depends crucially on superior long-distance seismological, radiological, and other sorts of monitoring. But in general, this is one scientific field that has more than fulfilled its promise. See Torrey Froscher, "Anticipating Nuclear Proliferation: Insights from the Past," *Nonproliferation Review* 13 (November 2006), p. 467–77.
14. Donald Mackenzie and Graham Spinardi, "Tacit Knowledge, Weapons Design, and the Uninvention of Nuclear Weapons," *American Journal of Sociology* 101 (July 1995), esp. pp. 89–91. Hugh Gusterson also makes this point, as well as noting other, more social functions that testing performs for the nuclear weapons laboratory design community. Hugh Gusterson, *Nuclear Rites: A Weapons Laboratory at the End of the Cold War* (Berkeley: University of California Press, 1996).
15. Ashley Tellis, *India's Emerging Nuclear Posture: Between Recessed Deterrent and Ready Arsenal* (Santa Monica, CA: RAND, 2001), pp. 198–99.
16. William Potter, "Remembering Nonproliferation Principles," in Jeffrey Laurenti and Carl Robichaud, eds., *Breaking the Nuclear Impasse: New Prospects for Security against Weapons Threats* (New York: Century Foundation Press, 2007), pp. 61–68.
17. Tellis, *India's Emerging Nuclear Posture*, p. 368.
18. Martin J. Sherwin, *A World Destroyed: Hiroshima and Its Legacies*, 3rd ed. (Stanford, CA: Stanford University Press, 2003).
19. Marc Trachtenberg, *History and Strategy* (Princeton: Princeton University Press, 1991), ch. 1.
20. Dan Lindley, "What I Learned Since I Stopped Worrying and Studied the Movie: A Teaching Guide to Stanley Kubrick's *Dr. Strangelove*," *PS: Political Science and Politics* 34 (September 2001), pp. 663–67.
21. The phrase is Jacques Derrida's. Bryan C. Taylor, "Pictures and Metapictures," *American Literary History* 9 (Autumn 1997), p. 567.
22. For more on this, see Jacques E.C. Hymans, "The Roots of the Washington Threat Consensus," in Betty Glad and Chris J. Dolan, eds., *Striking First: The Preventive War Doctrine and the Reshaping of U.S. Foreign Policy* (New York: Palgrave Macmillan, 2004), esp. pp. 41–42.
23. Albert Wohlstetter, Gregory Jones, and Roberta Wohlstetter, "Why the Rules Have Needed Changing," Part I of *Towards a New Consensus on Nuclear Technology*, vol. 1 (Los Angeles: Pan Heuristics, 1979), p. 36. Emphasis in original.
24. Ibid., p. 37.
25. Moreover, they also argue that the IAEA's traditional estimates of how much fissile material is needed to constitute an SQ are generally far too high. See Sokolski, ed., *Falling Behind,* esp. p. 25.
26. Avner Cohen, *Israel and the Bomb* (New York: Columbia University Press, 1998).
27. Note, however, that there is some indication that Israel may have engaged in testing jointly with South Africa during the 1970s. If Israel did cooperate on testing with South Africa, then there is no Israeli exception to the rule and the case for moving away from the test/no-test indicator more or less falls apart. On Israeli-South African cooperation see Purkitt and Burgess, *South Africa's Weapons of Mass Destruction*, pp. 51–52.
28. Most recently Prime Minister Ehud Olmert. See Philippe Naughton, "Olmert's Nuclear Slip-up Sparks Outrage in Israel," *Times Online*, December 12, 2006, <www.timesonline.co.uk/tol/news/world/middle_east/article752059.ece>.
29. Adcock and Collier, "Measurement Validity," p. 531.
30. Avner Cohen and Benjamin Frankel, "Opaque Nuclear Proliferation," in Frankel, ed., *Opaque Nuclear Proliferation: Methodological and Policy Implications* (London: Frank Cass, 1991), pp. 14–44.
31. See, e.g., Keith B. Payne, *Deterrence in the Second Nuclear Age* (Lexington, KY: University Press of Kentucky, 1996); Paul Bracken, *Fire in the East: The Rise of Asian Military Power and the Second Nuclear*

*Age* (New York: HarperCollins, 1999); Victor Cha, ''The Second Nuclear Age: Proliferation Pessimism versus Sober Optimism in South Asia and East Asia,'' *Journal of Strategic Studies* 24 (2001), pp. 79–120; Noah Feldman, ''Islam, Terror and the Second Nuclear Age,'' *New York Times Magazine*, October 29, 2006, <www.nytimes.com/2006/10/29/magazine/29islam.html?pagewanted=all>.

32. E.g., Shlomo Aronson with Oded Brosh, *The Politics and Strategy of Nuclear Weapons in the Middle East: Opacity, Theory, and Reality, 1960–1991—An Israeli Perspective* (Albany: State University of New York Press, 1992).
33. See Zeev Maoz, ''The Mixed Blessing of Israel's Nuclear Opacity,'' *International Security* 28 (Fall 2003), pp. 44–77.
34. Recent years have also seen a selective opening up of the Israeli archives on the subject, possibly the result of a strategy to normalize the Israeli bomb in the eyes of the world. For his part, Avner Cohen explicitly embraces this normalization as his political objective; see Cohen, *Israel and the Bomb*, pp. 339–49.
35. Though it must also be said that few want Israel to leave it.
36. Whether they were *ever* in it is a judgment call. In my view, among these four only Pakistan is a likely case of operational induction of nuclear weapons without prior testing. See Table 1.
37. Avner Cohen, ''Israel: A Sui Generis Proliferator,'' in Alagappa, ed., *The Long Shadow*, pp. 241–68. In his original work on ''opaque proliferation,'' Cohen had already made reference to *sui generis* aspects of the Israeli case, but only to strongly reject the idea that these aspects reduced the generalizability of the Israeli model. He now places much more stress on its unique character.
38. ''Documents Seemingly Counter US Intelligence Claims on North Korean Plutonium Production,'' Global Security Newswire, May 29, 2008, <gsn.nti.org/gsn/GSN_20080529_94E038D3.php>.
39. David Albright, ''North Korea's Alleged Large-Scale Enrichment Plant: Yet Another Questionable Extrapolation Based on Aluminum Tubes,'' Institute for Science and International Security press release, February 23, 2007, <isis-online.org/uploads/isis-reports/documents/DPRKenrichment22Feb.pdf>.
40. Barton Gellman, *Angler: The Cheney Vice Presidency* (New York: Penguin Press, 2008).
41. Dafna Linzer, ''Nuclear Capabilities May Elude Terrorists, Experts Say,'' *Washington Post*, December 29, 2004, p. A1.
42. For more on this point see Hymans, ''Assessing North Korean Nuclear Intentions and Capacities.''
43. Farah Stockman, ''US, North Korea Met on Nuclear Program,'' *Boston Globe*, May 19, 2005.
44. Jungmin Kang and Peter Hayes, ''Technical Analysis of the DPRK Nuclear Test,'' Nautilus Institute Policy Forum Online 06-89A, October 20, 2006, <www.nautilus.org/fora/security/0689HayesKang.html>.
45. The best the US intelligence community could do was to state that it was ''probably'' a nuclear explosion. <www.dni.gov/press_releases/20090615_release.pdf>.
46. George Soros, ''The Bubble of American Supremacy,'' *Atlantic Monthly*, December 2003.
47. E.g., Sokolski, ed., *Falling Behind*.
48. Hymans, *The Psychology of Nuclear Proliferation*, ch. 1.
49. Elizabeth Kier and Jonathan Mercer, ''Setting Precedents in Anarchy: Military Intervention and Weapons of Mass Destruction,'' *International Security* 20 (Spring 1996), pp. 77–106.
50. The importance of choosing obvious and salient focal points was first elaborated in Thomas Schelling, *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1960).
51. Richard Ned Lebow, *The Tragic Vision of Politics: Ethics, Interests, and Orders* (Cambridge: Cambridge University Press, 2003).
52. See, for example, the International Panel on Fissile Materials, <www.fissilematerials.org>.
53. Adcock and Collier, ''Measurement Validity,'' p. 542.
54. Richard Ned Lebow and Janice Gross Stein, *We All Lost the Cold War* (Princeton, NJ: Princeton University Press, 1994).
55. Robert Keohane, Gary King, and Sidney Verba, *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton, NJ: Princeton University Press, 1994).
56. Peter Liberman, ''The Rise and Fall of the South African Bomb,'' *International Security* 26 (Autumn 2001), pp. 45–86; Purkitt and Burgess, *South Africa's Weapons of Mass Destruction*.
57. See Michael E. Brown, Owen R. Coté Jr., Sean M. Lynn-Jones, and Steven E. Miller, *Offense, Defense, and War: An International Security Reader* (Cambridge, MA: MIT Press, 2004). For a dissenting view, see Elizabeth Kier, *Imagining War: French and British Military Doctrines Between the Wars* (Princeton: Princeton University Press, 1997).

58. Keir A. Lieber, *War and the Engineers: The Primacy of Politics over Technology* (Ithaca, NY: Cornell University Press, 2005).
59. See S. Paul Kapur, *Dangerous Deterrent: Nuclear Weapons Conflict and Deterrence in South Asia* (Stanford, CA: Stanford University Press, 2007).
60. Froscher, ''Anticipating Nuclear Proliferation.''
61. Purkitt and Burgess, *South Africa's Weapons of Mass Destruction*, p. 45.
62. For the internal U.S. government debates on the matter, see Jeffrey Richelson, ed., ''The Vela Incident: Nuclear Test or Meteoroid?'' National Security Archive Electronic Briefing Book No. 190, 2006 <www.gwu.edu/~nsarchiv/NSAEBB/NSAEBB190/index.htm>. These debates were mainly technical in nature. Based on their *political* analysis, Purkitt and Burgess in *South Africa's Weapons of Mass Destruction* conclude that there was indeed most likely a test (pp. 44–45).
63. Liberman, ''The Rise and Fall of the South African Bomb,'' p. 53.
64. Richelson, *Spying on the Bomb*, p. 346.
65. Ibid., chs. 6 and 8.