The Comprehensive Test Ban Treaty: Effectively Verifiable

In October 1999, the U.S. Senate declined to consent to ratification of the 1996 Comprehensive Test Ban Treaty (CTBT). Although the floor debate itself was so short as to be perfunctory, verification concerns played a role in the Senate’s action. At the time, some senators contended that the treaty’s verification provisions were inadequate to deal with potential cheating, despite assertions to the contrary by the White House.

Since then, the Bush administration has refused to reconsider the CTBT, despite widespread, almost unanimous international support for the pact. In the past year or two, however, there has been growing support across the U.S. political spectrum for Senate reconsideration of the CTBT. A bipartisan group of four senior statesmen, including former Secretaries of State George Shultz and Henry Kissinger, former Secretary of Defense William Perry, and former Senate Armed Services Committee Chairman Sam Nunn (D-Ga.), have called for the Senate to re-examine and ratify the treaty, which prohibits the conduct of any nuclear weapons test explosion or any other nuclear explosion anywhere.

The Democratic presidential nominee, Senator Barack Obama (Ill.), has also endorsed this call. In a July 16 speech in Indianapolis, he said, “[W]e’ll work with the Senate to ratify the Comprehensive Test Ban Treaty and then seek its earliest possible entry into force.” The Republican presidential nominee, Senator John McCain (Ariz.), who voted against the treaty in 1999, said in a May 27 speech that if elected, he would “begin a dialogue with our allies, and with the U.S. Senate, to identify ways we can move forward to limit testing in a verifiable manner that does not undermine the security or viability of our nuclear deterrent. This would include taking another look at the Comprehensive Test Ban Treaty to see what can be done to overcome the shortcomings that prevented it from entering into force.” How McCain might try to address perceived shortcomings on the verifiability of the CTBT is not clear.

The high-level calls for CTBT reconsideration reflect the growing recognition that renewed U.S. leadership on nuclear nonproliferation and disarmament is needed to shore up the nuclear Nonproliferation Treaty (NPT). The CTBT has long been seen as a litmus test of the nuclear-weapons states’ commitment to fulfilling their NPT Article VI commitment to end the arms race and pursue measures leading to nuclear disarmament. As of Sept. 26, 2008, 180 states have signed the CTBT and 145 have ratified, including all U.S. NATO allies. Iraq is the most recent signatory. However, the United States and eight other states, which
are necessary for CTBT entry into force, have not ratified the treaty.

Since 1996, only three nations have defied a de facto global nuclear test moratorium: India, North Korea, and Pakistan. In each case, the international community immediately and strongly condemned the tests. Moreover, international verification capabilities have been tested and found sufficient to verify compliance with the CTBT. A particularly noteworthy example occurred in October 2006 when the extensive network of sensors in the International Monitoring System (IMS) set up to monitor CTBT compliance easily detected a relatively low-yield (0.6 kiloton) nuclear test explosion by North Korea. This detection occurred even without the full spectrum of nuclear explosion verification and monitoring capabilities that would be provided for under the CTBT. These capabilities include the possibility of short-notice, on-site inspections.

When and if U.S. policymakers decide to reconsider the CTBT, they will find that CTBT entry into force would not only strengthen the norm against testing, but would increase the technical and legal capabilities to monitor, detect, and deter violations of the test ban norm. They will also discover that since the 1999 Senate vote on the CTBT, there have been significant improvements in technical verification capabilities. Together, these developments mean that few questions should remain about whether the CTBT meets the standard for “effective verification” as set down by top national security officials when the Senate considered and eventually supported previous arms control treaties.

**How Much Verification Is Enough for Effective Verification?**

In his Sept. 22, 1997, letter transmitting the CTBT to the Senate, President Bill Clinton wrote that “our National Intelligence Means, together with the Treaty’s verification regime and our diplomatic efforts, provide the United States with the means to make the CTBT effectively verifiable.”

A second attempt to convince a two-thirds majority of the Senate that the CTBT is verifiable will depend in part on achieving a clear understanding of what effective verification means. The U.S. standard for effective verification of an arms control treaty was defined during Senate consideration of ratification of the 1988 Intermediate-Range Nuclear Forces (INF) Treaty and the 1991 START. During INF Treaty ratification hearings, Ambassador Paul Nitze defined effective verification: “If the other side moves beyond the limits of the treaty in any militarily significant way, we would be able to detect such violation in time to respond effectively and thereby deny the other side the benefit of the violation.”

Thus, cheating that could threaten U.S. national security in a militarily significant way must be detected in sufficient time. In the case of a nation that already has nuclear weapons, effective verification is determined by the military significance of the additional nuclear-weapons capabilities it might obtain by cheating, beyond those it had before the treaty was in place.

During the START ratification hearings, Secretary of State James Baker repeated the Nitze definition of effective verification but added a criterion: “Additionally, the verification regime should enable us to detect patterns of marginal violations that do not present immediate risk to U.S. security.”

The intelligence community has a somewhat different approach based on statistical criteria. High-confidence verification, in the community’s view, is the ability to detect 90 percent of the violations of treaty provisions. For the CTBT, this means the ability to detect 90 percent of nuclear tests

---

**Figure 1: Explosion Detection by the International Monitoring System**

A 2002 National Academy of Sciences (NAS) study estimated the ability of the International Monitoring System (IMS) of the Comprehensive Test Ban Treaty Organization (CTBTO) to detect a clandestine nuclear explosion. The NAS study concluded that nuclear tests of well below one kiloton in explosive yield would be detected by the IMS as shown in the figure. Depicted below are contours indicating the yields of tamped explosions that would be detected at various locations by three or more IMS seismic stations. Yields are expressed in tons (t) (60 tons is equivalent to 0.06 kilotons).
Figure 2: Identifying Low Yield Explosions-The North Korean Example

Seismograms of the October 9, 2006 North Korean test show that nuclear explosions can be differentiated from earthquakes. The graphic below depicts seismograms recorded in northeast China from the North Korea nuclear test (top); from an earlier nearby earthquake (middle); and from a small chemical explosion (bottom). The nuclear test (magnitude 4, approximately 0.8 kilotons) can be distinguished from the earthquake (magnitude 4) by the fact that the nuclear test has stronger primary waves and weaker shear waves. The primary wave amplitude oscillates in the direction the wave travels, whereas the shear wave amplitude oscillates at right angles to the travel direction, such as a “wave” at a sporting event where people oscillate up and down but the wave travels horizontally. The nuclear test can be differentiated from a two-ton chemical explosion (magnitude 1.9) with a similar pattern of primary and shear waves by the fact that the chemical explosion’s seismic signal at this station is of lower quality and is about fifty times smaller than the nuclear test. Combining seismic data with other tests, such as air sampling for radionuclides, indicates that nuclear explosions below 0.010 kiloton yield (below 10 tons yield) could be detected in the region.

CTBT Monitoring Progress Since the 1999 Senate Vote

The 1963 Limited Test Ban Treaty bans nuclear testing in the atmosphere, outer space, and underwater and is verified solely by classified national means. In contrast, the CTBT bans nuclear testing everywhere, and it has extensive verification provisions. These include the IMS, consisting of remote sensors (seismic, radionuclide, hydroacoustic, and infrasound), confidence-building measures; provisions for consultation and clarification; and, once the treaty enters into force, the possibility of short-notice, on-site inspections, as well as national means. CTBT verification concerns have primarily focused on possible cheating by underground testing.

When the Senate considered the CTBT, some treaty critics suggested that it could only be monitored at testing levels of about one kiloton. This miscalculation arose in part from the fact that CTBT negotiators decided that the IMS should be designed to monitor nonnuclear detonations above a threshold of about one kiloton anywhere in the world with high confidence. In reality, however, actual test ban verification capabilities were much better and have only improved since then. Not only have IMS monitoring technologies gotten better, but so have the capabilities of civilian seismic networks and national technical monitoring systems.

Indeed, a 2002 panel of the National Academy of Sciences (NAS) determined that “underground nuclear explosions can be reliably detected and can be identified as explosions, using IMS data down to a yield of 0.1 kilotons (100 tons) in hard rock if conducted anywhere in Europe, Asia, North Africa and North America” (see fig. 1). Advances in regional seismology have provided additional confidence. For some locations, such as Russia’s former nuclear test site at Novaya Zemlya, the use of new seismic arrays and regionally located seismic stations has lowered the threshold to below 0.01 kilotons.

As an example, take the 0.6-kiloton, North Korean test noted above. This nuclear explosion was promptly detected and identified from signals recorded at 31 seismic stations in Asia, Australia, Europe, and North America, including 22 IMS stations established by the Preparatory Commission for the Comprehensive Test Ban Treaty Organization. The U.S. Geological Survey posted a good estimate of the test location five hours after the explosion occurred.

Seismic data from underground chemical explosions show that far-lower-yield explosions would have been detected in the Korean region, as low as 0.002 kilotons in fact, a factor of 50 below the 0.1-kiloton NAS monitoring threshold and a factor of 500 below the nominal 1-kiloton threshold (see fig. 2). The 0.002-kiloton threshold is not applicable everywhere, but it is evidence that considerably lower limits can be obtained than previously assumed, particularly if there is a sufficient density of regional seismic stations.

Regional monitoring is based on signals that have traveled via the earth’s crust and upper mantle and have been recorded at distances up to about 1,500 kilometers. Better results are obtained with regional monitoring than with longer-range or teleseismic stations, which measure body waves that travel below the mantle. New algorithms, closer access, and seismic models enhance the ability to improve location...
estimates and discriminate better among nuclear tests, earthquakes, chemical explosions for mining, or other phenomena. For example, scientists have examined whether regional data from seismographs located at a distance of 500-1,500 kilometers could shed new light on 69 (out of 340) Soviet underground nuclear tests that took place at the Semipalatinsk Test Site in Kazakhstan during the Cold War but whose origin times, coordinates, and magnitudes had not been publicly determined. Using regional data, scientists were able to provide information on all but two tests over 1 ton (0.001 kilotons). This achievement took place with seismographs employing old technology. Newer technology provides even more potential.10

Moreover, regional monitoring has been enhanced by the continued global expansion of the IMS. As of Sept. 26, 2008, 233 of the 337 IMS facilities were certified, 30 were operational but not yet certified, 36 were under construction, and 38 are planned. Since 89 percent of the IMS facilities are now certified, operational, or under construction, it is reasonable to expect that 95 percent of the IMS network will be completed in about five years. Additional data can be retrieved quickly from the 120-station Auxiliary IMS network and from the vast Global Seismic Network.

The IMS has also improved the ability to detect radionuclides that would be emitted after a nuclear explosion by an order of magnitude. Forty of 80 radionuclide stations will be able to monitor for particulate and gaseous radionuclides. Radionuclides from the subkiloton North Korean test were detected in nearby South Korea and in distant Yellowknife, Canada, more than 7,000 kilometers away.

Experts have also been able to lower thresholds by comparing waveforms from two separate seismic events that are in close proximity. The small relative differences in the waveforms are used to reduce location uncertainties in seismically active regions by factors of 10 to 100, as compared to estimates based on arrival times of seismic waves. This is extremely useful at former test sites, and the technique is being extended to earthquake regions.

A new detection technique, interferometric synthetic aperture radar (InSAR), is widely used by the United States, with its four Lacrosse satellites, as well as by the European Space agency, Canada, and Japan. InSAR can measure the earth's subsidence (in many cases, to 2-5 millimeters accuracy) after a disruption, depending on the local situation.11 InSAR is not used to measure nuclear yield but rather to pinpoint the location of the explosion to within 100 meters. In addition, InSAR can discriminate between earthquakes and explosions on the basis of the symmetry of the subsidence pattern. The detection threshold for InSAR is less than 1 kiloton at 500 meters depth for many locations if prior radar data is available, which is now frequently the case. Seismic data of a suspicious event can direct InSAR measurements to enhance the interpretation of the seismic data, and InSAR can provide much better location accuracy to direct on-site inspections of an ambiguous event. Its accuracy of 0.01 square kilometers is vastly better than the CTBT's allowed upper limit of 1,000 square kilometers.

Evasion Scenarios
In its 2002 report, the NAS panel examined 10 evasion scenarios suggested by the U.S. intelligence community, concluding that the only evasion scenarios that need to be taken seriously at this time are cavity decoupling and mine masking.12 The most commonly cited concern during the Senate debate was cavity decoupling, which is the use of a cavity to muffle the seismic wave from a nuclear explosion. No country is known to have fully decoupled a nuclear explosion in a cavity that was created for that purpose. An explosion is fully decoupled if the cavity is large enough to reduce the nuclear blast pressure on the cavity wall below a critical level (the cavity radius to do this must be larger than 25 meters multiplied by the yield in kilotons to the 1/3 power). The only fully decoupled nuclear test to do this displayed a yield that was reduced by a factor of 70. This was the 1966 Sterling test with a yield of 0.38 kilotons in a 34-meter diameter Mississippi salt cavity, which had been produced from an earlier 5.3-kiloton nuclear explosion. In 1976 the Soviet Union managed to partially decouple by only a factor of 12 with a test in a salt cavity in Azgir. The 9-kiloton weapon was too large to fully decouple in the 72-meter diameter cavity, which had been created by an earlier 64-kiloton explosion.

During the 1999 Senate debate, then-Senate Majority Leader Trent Lott (R-Miss.) mistakenly claimed that a 70-kiloton explosion in a cavity could be hidden from IMS monitoring. To do this would require a cavity with a 200-meter diameter (equivalent to a 50-story building) with an area of 0.13 square kilometers at a depth of 1 kilometer. No such cavity has ever been constructed, and it would be essentially impossible to construct one physically or to use such a hypothetical cavity with sufficient secrecy to hide a test.

The NAS panel determined that an explosion in a cavity cannot be confidently hidden if its yield is larger than 1 or 2 kilotons. The prospect that a country could cheat undetected fails to take into account that nowadays arrays of seismographs and other seismic capabilities can detect and identify many events that take place more than 2,000 kilometers away with yields considerably less than 1 kiloton. It also ignores advances in regional seismology (discussed below). One must also factor in the fact that, for example, 90 percent of underground Soviet tests at Novaya Zemlya vented radioactive debris, as did 40 percent of all Soviet tests.13 Moreover, one must take into account that, for a successful clandestine test, the actual yield must be limited to the level predicted rather than a higher level (a "yield excursion"). This is a particularly difficult task for new nuclear-weapons states. The panel noted that if an inexperienced state wanted to reduce the risk of detection, "it would probably try to limit test yields to 0.1 [kiloton] or less."

Venting of radioactive gases is also a problem for any state that might consider a clandestine nuclear test. In fact, it is highly likely that a significant amount (more than 0.1 percent) of radionuclides would be released and then detected by the IMS. The congressional Office of Technology Assessment reported in 1989, for example, that "since 1970, 126 [U.S.] tests have resulted in radioactive materials reaching the atmosphere with a total release of about 54,000 curies. Of this amount, 11,500 curies, roughly one-fifth were due to containment failure and late-time seeps.14 Venting from smaller tests can be even more difficult to contain, as the last four U.S. tests that vented had yields of less than 20 kilotons. Some scientists hypothesize that smaller explosions may not sufficiently enclose cavities with a glassified cage to prevent venting. Furthermore, the cavities may not rebound sufficiently to seal fractures with a stress "cage," making venting more likely and the possibility of avoiding detection even smaller.

In the 1999 Senate debate, treaty op-
Few questions should remain about whether the CTBT meets the standard for “effective verification” as set down by top national security officials when the Senate considered and eventually supported previous arms control treaties.

these tasks could be carried out with high confidence of avoiding detection, there would only be a cumulative 50 percent chance of avoiding detection of one test and only a 15 percent chance that three tests could be carried out without detection. Yet, it is unlikely that a new nuclear-weapon state even could do this well given the major technical hurdles.

- Violators must be able to avoid yield excursions to prevent a planned clandestine test from being easily detectable. All first tests, including the early atmospheric tests if conducted underground, would have been detectable by the IMS: United States (1945, 21 kilotons), Soviet Union (1949, 20 kilotons), United Kingdom (1952, 25 kilotons), France (1960, 65 kilotons), China (1964, 22 kilotons), India (1974, 12 kilotons), Pakistan (1998, 9 kilotons), North Korea (2006, 0.6 kilotons).

- It is necessary to conceal from satellites materials removed to create a test shaft and cavity.

- Crater and surface changes due to testing must be hidden from InSAR and other technologies. The 1998 Indian and Pakistani test sites were easily located in SPOT commercial images with 5-meter resolution. As noted above, surface subsidence of a few millimeters is observable in many locations with InSAR.

- Practically all the radioactive gases and particles must be trapped. Radiocnuclide venting is a particularly serious risk for beginning nuclear-weapon states.

- The cheater must avoid the detection of weaker seismic signals by regional seismic stations and seismic arrays. With the advent of more seismic stations in the world, this becomes more difficult.

- The cheater must prevent the detection through national means. Human and other intelligence can provide other information that reveals test preparations. This information can be used by the CTBT Executive Council to authorize an on-site inspection.

What Kind of Cheating Would Matter?
It should be kept in mind that the principal risk that needs to be avoided is that a country under the CTBT could alter the strategic balance between it and the United States.

As noted above, the NAS study concluded that it would be very difficult for states with less nuclear testing experience, such as India, Iran, Iraq, North Korea, and Pakistan, to meet the required conditions to avoid detection by testing at 1 kiloton or less. It is far easier to test in the 10-kiloton range without a specific yield than to test at a specific small yield. The NAS study concluded that “[c]ountries of lesser prior test experience and/or design sophistication” would also lack the sophisticated test-related expertise to obtain “limited improvement of efficiency and weight of unboosted fission weapons compared to 1st-generation weapons not needing test-

ing” from tests at levels of 0.01 kiloton to 1-2 kilotons.

The NAS panel also judged that “[s]tates with extensive prior test experience [Russia and China] are the ones most likely to be able to get away with any substantial degree of clandestine testing.” For example, with difficulty, these states could validate designs for unboosted fission 1-kiloton weapons in a cavity. Yet, very-low-yield tests by nuclear-weapon states, such as Russia and China, should not materially change the strategic balance by themselves. A 1995 JASON panel concluded that testing at 0.5 kilotons would provide only minimal gains in developing a new weapon design. Moreover, at a minimum, several clandestine tests are needed to change design parameters, improving the chance of detection.

Net Benefit Analysis of Three CTBT Situations
The NAS panel examined three global situations: with full compliance to the CTBT, without the CTBT, and with the CTBT with evasions by China, India, Iran, Iraq, North Korea, Pakistan, and Russia. The NAS panel concluded that “[t]he worst-case scenario under a no-CTBT regime poses far bigger threats to U.S. security interests—sophisticated nuclear weapons in the hands of many more adversaries—than the worst-case scenario of clandestine testing in a CTBT regime, within the constraints posed by the monitoring system.”

Our discussion has covered only technical issues in a world that is much more complex in its behavior. General John Shalikashvili, former chairman of the Joint Chiefs of Staff, examined the net benefit of the CTBT by examining all aspects, including political ramifications for two worlds, with and without a CTBT. Shalikashvili’s report suggested a mechanism for CTBT ratification in which the United States would “commit to conducting an intensive joint review of the Test Ban Treaty’s net value for national security ten years after U.S. ratification, and at ten-year intervals thereafter. If, after these steps, grave doubts remain about the Treaty’s net value for U.S. national security, the President, in consultation with Congress, would be prepared to withdraw from the Test Ban Treaty under the ‘supreme national interests’ clause.” It is widely believed that the United States will not test a nuclear
weapon because China and Russia would quickly respond with multiple nuclear tests, leading to a new arms race. In addition, the prevailing wisdom in Washington is that the U.S. effort to halt nuclear proliferation would be greatly damaged if the United States tested. Thus, the United States is constrained by the CTBT, while not gaining the full benefits that would follow from the CTBT being in force.

The CTBT is Effectively Verifiable
Table 1 displays the monitoring limits for the various CTBT sensor systems from the 2002 NAS report and from other sources. As discussed above, cheating under the CTBT would be far from simple. Even if a country tested in a cavity and managed to carry out each of the six tasks identified above in such a way that the states might have high confidence (90 percent certainty for each task) that each task would not be detected, the cumulative difficulty of hiding such a test would be considerable and almost insurmountable for more than one test.

Still, to provide further assurance, additional monitoring capabilities could be deployed at declared test sites to address extremely small tests, but only after CTBT ratification and then to be followed by fur-

<table>
<thead>
<tr>
<th>Table 1: Tools for Comprehensive Test Ban Treaty Verification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technology/Method</strong></td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td><strong>Seismic</strong></td>
</tr>
<tr>
<td><strong>Hydroacoustic</strong></td>
</tr>
<tr>
<td><strong>Infrasound</strong></td>
</tr>
<tr>
<td><strong>Radionuclide</strong></td>
</tr>
<tr>
<td><strong>Interferometric Synthetic Aperture Radar (InSAR)</strong></td>
</tr>
<tr>
<td><strong>On-Site inspections</strong></td>
</tr>
<tr>
<td><strong>Confidence-Building Measures</strong></td>
</tr>
</tbody>
</table>

The Provisional Technical Secretariat of the Comprehensive Test Ban Treaty Organization is responsible for establishing an International Monitoring System (IMS) to assist with verifying compliance with the CTBT. The IMS will span the globe with 321 monitoring stations consisting of four technologies (seismic, radionuclide, hydroacoustic, and infrasound). It includes the infrasound station, shown above, under installation in Kersa, Tunisia.

ENDNOTES

1. In annual votes in 2003 to 2007, the UN General Assembly (UNGA) cumulatively gave 870 votes in favor of the treaty and only seven against (North Korea and Palau each, and the United States five times). By year, tally, and resolution, the votes were in 2007 (176/1/4, UNGA 62/59); 2006 (172/2/4, UNGA 61/104); 2005 (172/1/4, UNGA 60/95); 2004 (172/2/4, 59/109); and 2003 (173/1/4, UNGA 58/77). The 20 abstentions resulted from five votes each by Colombia, India, Mauritius, and Syria. Since Colombia ratified the CTBT on January 29, 2008, the number of total abstentions is effectively reduced to 15.


6. Article IV (D), paragraph 37 of the CTBT allows for the consideration of evidence and data regarding possible violations from national technical means.

7. NAS CTBT study, p. 57. The limit of 0.1 kiloton does not consider evasive measures.


9. NAS CTBT study, p. 56.


12. NAS CTBT study (2003), pp. 48-49.


15. NAS CTBT study, p. 68.


17. NAS-CTBT study, pp. 70-78.


19. Senate Committee on Foreign Relations, START Treaty, pp. 49-64; Senate Committee on Foreign Relations, START II Treaty, 104th Cong., 2d sess., pp. 29-37. The committee ratification reports to the Senate concluded that, because of the resilience of the U.S. nuclear triad, the surviving U.S. strategic forces would not be greatly affected by Russian cheating in yields or numbers of strategic weapons.