

## SEISMOLOGICAL METHODS FOR MONITORING A CTBT: THE TECHNICAL ISSUES ARISING IN EARLY NEGOTIATIONS

PAUL G. RICHARDS<sup>1</sup> and JOHN ZAVALLES<sup>2</sup>

<sup>1</sup>*Lamont-Doherty Earth Observatory, and Department of  
Geological Sciences, Columbia University,  
Palisades, New York 10964, USA.*

<sup>2</sup>*US Department of Defense, Office of Humanitarian and Refugee  
Affairs*

### 1. Introduction

The most important technical issues in monitoring a Comprehensive Test Ban Treaty all became apparent between 1958, when the so-called “Conference of Experts” was convened in Geneva, and 1963, when the Limited Test Ban Treaty (LTBT) was negotiated, signed, ratified, and put into effect. The LTBT was negotiated trilaterally, between the United States, the Soviet Union, and the United Kingdom; and banned nuclear testing in space, in the atmosphere, and under water. But in 1963 there was a general perception that seismological methods for monitoring underground nuclear explosions were inadequate – a perception that helped strongly to prevent the conclusion of a Comprehensive Test Ban Treaty (CTBT) in this period, even though the key leaders (Presidents Eisenhower and Kennedy for the US, General Secretary Khrushchev for the USSR, and Prime Minister Macmillan for the UK) apparently favored a comprehensive ban and made numerous proposals on how it might be implemented and verified.

In the words of a former Director of the Los Alamos National Laboratory, Dr. Donald M. Kerr, “Nuclear weapon testing is ... a process intimately intertwined with the design of nuclear weapon systems” [1]. In the three decades following the unsuccessful CTBT negotiations of the early 1960s, over 1500 underground nuclear tests were carried out by the US, the USSR, France, the UK, China, and India; that is, about one nuclear test a week, for thirty years. Though the great majority were single explosions, more than a hundred of the tests consisted of more than one nuclear explosion. It appears from this extensive activity that nuclear weapons development, at least by the

superpowers, has not been constrained by the LTBT.

Very few underground tests had been carried out, at the time of the early CTBT negotiations. Indeed, only one such test had occurred prior to the 1958 Geneva meetings. The early conclusions on the capability of seismological methods to monitor a CTBT were therefore reached on the basis of minimal data, at a time when seismology itself was practiced at only a few tens of institutions around the world. Seismographic data then consisted usually of paper records, rarely seen outside the institution that owned them; were derived entirely from earthquakes or small chemical explosions; covered only a narrow range of frequencies; and had low dynamic range so that it was not possible to record both strong and weak signals on the same instrument.

The perceived failure of seismology to support a major arms control objective – the CTBT – coupled with the need to acquire information on foreign underground nuclear weapons tests, led the US and the USSR in about 1960 to begin new efforts in instrumentation and research. These efforts hastened the development of seismology over more than a decade, turning it into a modern quantitative science.

This paper first reviews the key technical issues in CTBT verification that emerged in politically charged negotiations up to 1963, and comments upon the conclusions concerning monitoring capability reached by the US at that time. In a later section, the key technical issue not developed prior to 1963 (namely, event identification) is briefly reviewed in light of later practical experience. The underlying question throughout is not only: “How are underground nuclear explosions detected and identified?” We also ask: “How did our present understanding of these issues evolve?” By the late 1960s, seismological monitoring methods had improved to reach the level of CTBT verification capability desired by the U.S. in 1963, even though the global monitoring network advocated in 1958-60 was never built. Had it been built, monitoring capability would have been improved about tenfold over what was desired by the U.S. in 1963.

## **2. History of the Treatment of CTBT Verification, 1957-1963**

The failure of post-World War II efforts to develop some type of international control over atomic weapons, and the subsequent development of the hydrogen bomb, resulted in substantial programs of atmospheric nuclear weapons testing. In 1954, fallout from the US 15 megaton BRAVO test contaminated a Japanese fishing boat, causing the death of one man and the serious illness of several others. Later that year, radioactive debris from a Soviet test fell over part of Japan. Many physicians and biologists, now able to observe the long-term medical effects of the Hiroshima and Nagasaki bombings, charged that atmospheric tests would carry radioactive material worldwide, causing a genetic hazard that would be particularly damaging if cumulative doses occurred from fallout.

The dangers of fallout thus became the first rationale for a test ban, with proponents such as Linus Pauling, the 1954 Nobel Prize winner in chemistry, predicting that more people would die of cancer if atmospheric testing continued unchecked; and opponents such as Edward Teller, a promoter of nuclear weapons development, concluding that cigarette smoking would shorten lives far more than fallout. At this time Edward Teller and Ernest Lawrence, of the Livermore Radiation Laboratory, sought to separate the fallout and testing issues by speaking of a “95% clean” hydrogen bomb, whose fallout would be negligible. However, such a device had to be detonated very high in the atmosphere to minimize fallout, and had to be of very high yield, relative to the size of its fission trigger, to approach the 95% claim [2]. In February 1955 the US Atomic Energy Commission published a report on fallout and its consequences, which was widely criticized as a biased justification of nuclear testing [3, p. 140].

By the mid-1950s both superpowers realized the need for discussion, at least, of a prohibition on nuclear weapons testing. However, they differed significantly on how a test ban might be implemented. Initially the Soviets favored an agreement to outlaw tests, preceded by a complete testing moratorium, before a control system to monitor the agreement was established. Because such a system could entail intrusive on-site inspections, they claimed that inspections with no prior treaty commitment would aid US espionage efforts and undermine Soviet security. The United States preferred the establishment of formal controls prior to the signing of an agreement. It was believed in the West that without such controls the more restrictive nature of Soviet society might permit clandestine testing, and consequently a lead over the US in weapon development. However, at the 1957 meeting of a Subcommittee of the Disarmament Commission of the UN, conducted in London, the USSR surprisingly announced that it would agree to establishing a control system with posts in the USSR, US, UK, and somewhere in the Pacific Ocean, prior to the signing of a test ban agreement. Furthermore, the Soviets announced they would accept a temporary moratorium on tests, two to three years in length. Most importantly, the Soviets accepted a British suggestion that the Subcommittee establish technical working groups to study the feasibility of limiting tests and monitoring such an agreement. At this time, it should be noted, the USSR proposed a nuclear test ban as an objective unto itself, while the Western powers favored a test ban only as a first step to some type of general disarmament agreement [4, p. 16].

Two events, occurring far from the negotiating table, rendered the test ban debate more urgent. Thus, on September 19, 1957, the US carried out the world’s first underground nuclear explosion in which the radioactive by-products were fully contained. This was the 1.7 kiloton RAINIER test on the Nevada Test Site. Quoting from Dr. Gerald Johnson [5], writing of his experience when in charge of the nuclear test program for Livermore: “The development of the technique of underground testing was stimulated

in the latter 1950s because of rising concerns about radioactive fallout both locally and worldwide from atmospheric testing. These concerns were brought forcibly to my attention in 1956 ... when we experienced delays of up to three weeks awaiting favorable wind patterns which would result in acceptable local fallout.” Second, on October 4, 1957, the Soviet Union launched the world’s first artificial satellite, Sputnik I, an event which galvanized the US scientific community. In November, Eisenhower established the President’s Science Advisory Committee, under the leadership of James Killian of MIT. The PSAC contained many men who opposed the unlimited testing and development posture of the Department of Defense and the Atomic Energy Commission. Foremost among these was Hans Bethe. During the difficult Geneva negotiations of the following year, the PSAC was to provide Eisenhower with views drastically different from those of Teller and other prominent US advisers.

Early in 1958 Killian appointed an inter-agency committee, including representatives of the PSAC, AEC, and DOD, to study the technical feasibility of monitoring a test ban. This panel, chaired by Hans Bethe, reported in April that a system of 24 inspection stations in the USSR, supplemented by mobile inspection teams, could detect underground explosions down to a yield of one or two kilotons. The panel also concluded that the risks to the US in being subjected to such a test ban would be small, while the political advantages were large; and that continued nuclear testing by both powers would rapidly erode the technological lead the US enjoyed over the USSR.

Note that the only technical data at the disposal of the Bethe Panel was from the RAINIER test conducted six months earlier. Note too that although judgments were made on the political ramifications of a test ban, no political scientists or diplomats were members of the Panel. The problem, which persists in test ban debates to this day, was that technical experts were called upon to address issues which were essentially political in nature. The difficulty arises, when the evaluation of technical capabilities (and associated uncertainties) turns into a review of whether such capabilities are in some sense acceptable — which is ultimately a political question. Test ban opponents were quick to highlight these shortcomings of the Bethe Panel, and Teller also emphasized that the Panel had completely ignored the study of intentional evasion by one side.

The level of technical discussion of the RAINIER data reached a low point when the AEC publicly announced that seismic signals from this shot were detected to a maximum distance of “only 250 miles” (400 km). After an outcry from knowledgeable scientists, the detection estimate was revised to 3700 km because of an observation in Alaska. However, inspection of the seismogram in question shows that during a 24-hour period it contained numerous “detections” with amplitude comparable to that from the RAINIER explosion. Today, we would speak of this as a problem not just in *detection*, but in *association*. Given the level of seismic activity around the globe (with approximately ten earthquakes a day having seismic waves comparable in size to

those from RAINIER), and the numerous instruments now deployed to record seismic signals, it is necessary not only to detect signals, but to form the correct subset of detections from different stations for a common seismic event – whether earthquake or explosion – before proceeding to analyse the set of detections, for example to estimate the location of the source.

The controversy over the Bethe Panel's findings coincided with an announcement of the Supreme Soviet on March 31, 1958, that the USSR would discontinue all nuclear tests providing the US and the UK followed suit. Nine days before, the Soviets had concluded a very extensive test series, in the course of which two or three explosions were often conducted in a single day. The US was awkwardly placed, as it was scheduled to commence its own test series in several weeks. Sensing a public relations ploy, Eisenhower dismissed the Soviet announcement as a "gimmick," which ought not to be seriously considered [6].

Letters ensued between Eisenhower and General Secretary Khrushchev, and on April 8 Eisenhower proposed a meeting of technical experts, as envisioned at the London Conference, a proposal which Khrushchev initially rejected on the grounds that the conclusion of a test ban was a political, not a scientific, matter. Eisenhower reiterated the need for technical meetings on April 28, in a statement of some import: "Studies of this kind are the necessary preliminaries to putting political decisions actually into effect" [4, p. 50]. Some confusion existed in the interpretation of this sentence. The Soviets believed it defined the technical meetings as formalities, subsidiary to the inevitable conclusion of the treaty. The US interpretation was that the technical talks were meant to explore the feasibility of concluding a treaty in the first place.

Khrushchev accepted the scheduling of these talks on May 9, and the Conference of Experts (to Study the Possibility of Detecting Violations of a Possible Agreement on the Suspension of Nuclear Tests) convened on July 1, 1958 in Geneva. Although conducted at United Nations facilities, the Conference was not a UN-sponsored activity.

The American delegation consisted of James Fisk (Chairman), Robert Bacher, and Ernest O. Lawrence (who returned to the US due to illness in the middle of the conference and died shortly thereafter). A group of about a dozen physicists and seismologists was on hand to advise the delegates. The only US State Department representatives were three observers of relatively junior rank, whereas the Soviet delegation, headed by Yevgeni Federov, included one very high ranking diplomat, Semyen Tsarapkin of the Collegium of the Ministry of Foreign Affairs. The difference in composition of the panels highlighted contrasting attitudes regarding the Conference and its eventual goals.

The Conference first discussed detection and identification in four media: the surface and atmosphere of the Earth; underwater; in space; and underground. Panelists generally agreed that tests in the first category could be readily detected by their output of acoustic and radio waves and radioactive debris, and oceanic tests could be easily detected with

hydroacoustic waves. It was agreed that when an underground test is conducted at a depth sufficient to prevent radioactive debris from reaching the surface, signals produced by seismic waves were the only means of detection.

Although there are several different seismic waves, in these early years only the fastest seismic wave was given detailed consideration. This wave, known as the *P*-wave (the “*P*” standing for “*primus*,” since it is the first wave to arrive at a distant station), spreads though the Earth’s deep interior in much the same fashion that a pulse of sound moves through air. In air, the source of the sound might be an exploding firecracker. In the Earth, the source of *P*-waves might be an earthquake or an underground nuclear explosion. If enough instruments at different locations record the arrival time of the *P*-wave from a particular source, it then becomes possible to estimate the source location. Throughout the five years leading up to the LTBT of 1963, it appears with few exceptions that only *P*-waves were considered for use in event identification.

The experts in 1958 found few problems with detection by electromagnetic and hydroacoustic waves. A plan for airborne collection of radioactive debris was eventually worked out, in which aircraft belonging to the nation being overflown would be used, representatives of all the signatories would be on board, and the flight path would be determined in advance.

The issue of underground detection was far more complex. The only empirical data available at this time was for the RAINIER test. The Soviet delegation was in general optimistic, and held that theoretical interpretation of data from TNT explosions would be sufficient to determine monitoring capabilities for underground nuclear explosions. The Americans replied that the RAINIER data were not consistent enough to allow useful extrapolation.

After discussing detection methods, the Conference next turned to the monitoring system necessary. Federov immediately proposed a system of 100 to 110 control posts, the spacing of which would be based on acoustic detection of atmospheric tests. The Soviets maintained that the existing network of seismic stations for earthquake identification would suffice to monitor underground tests.

The British and Americans rejected this proposal, arguing that existing seismic stations were not adequate for the task and could serve only as a supplement to a future system of new seismic station or “control posts.” The only criterion for distinguishing earthquakes from explosions accepted at this time was the “first motion determinant,” which claimed that all explosions would be accompanied by a positive first motion (compression), resulting in *upward* movement of the surface of the ground at all distant monitoring stations when the *P*-waves arrived, while earthquakes would feature negative first motion (rarefaction), or *downward* movement of the ground, at least at some stations. The Western experts argued that in most stations the switching of a few wires

would reverse the recorded polarity of the first motion and make a compression appear to be a rarefaction. Thus, relying on many stations manned by the potential violator nation would be unacceptable.

It was recognized that the size of the seismic monitoring system was inextricably linked to the desired threshold, above which events could be not only detected but identified as an earthquake or an explosion. The Western delegates estimated that seismic systems already in place could detect and identify 5% of events equivalent to a one kiloton yield and up, 50% of 5 kt and above, and 90% of 20 kt and above [7, p. 26].

The job of developing a US counterproposal on the size of the monitoring system was entrusted to two young physicists, Richard Latter of the RAND Corporation and Harold Brown of the Livermore Laboratory. The criterion they proposed was observation of first motion at five different seismic stations. Advocating a capability of detecting and identifying 90% of events equivalent to explosions of one kiloton and up, they concluded that a global network of some 650 stations would be necessary. Although these findings were presented merely as a report, not as a formal proposal, Federov indicated that such a control system would be unacceptable to the USSR.

A third system was proposed by Sir William Penney, head of the British delegation. In its detection capability Penney's system was a compromise between the US and Soviet models. He suggested 160 to 170 land-based control posts, each operating a small array of about 10 seismic stations of which 100 to 110 would be based on continents, 20 on large islands, and 40 on small islands, although the precise locations were not specified. These posts would be supplemented by ten ships equipped to detect atmospheric and oceanic tests. Penney believed [7, pp. 26-35] that such a system would detect and identify 90% of the earthquakes equivalent to 5 kilotons or more, and a small percentage of those equivalent to one kiloton. The other 10% of 5 kt equivalent events would have to be inspected, and estimates of the number of such events ranged from 20 per year (USSR estimate) to 100 per year (US estimate).

The Penney system was in general well received, and the Conference concluded on August 21 after both groups agreed to recommend a system based on Penney's proposal. One important issue was left unresolved. The West suggested that all suspicious unidentified events lead to on-site inspections, while the Soviets favored individual decisions on each case by the control commission, with each member nation having veto power.

Many observers believed that, by obtaining agreement on Penney's proposal, the USSR had maneuvered the US into agreeing to a test ban. Critics such as Teller contended that the limitations of the system were not strongly presented, and the idea of deliberate evasion was not discussed. "In the contest between the bootlegger and the police," Teller wrote, "the bootlegger has a great advantage" [8]. He and many other critics believed the Conference of Experts constituted a Soviet propaganda victory [9].

Nonetheless, the US government, especially the State Department, felt compelled to make a public statement demonstrating its commitment to eventual test cessation. President Eisenhower proposed on August 22, 1958, that formal negotiations begin on October 31, and announced that the US would refrain from testing for a period of one year from the start of these talks unless the Soviets resumed tests. Such a moratorium could be extended, he went on to say, subject to the installation of a control system and satisfactory progress in implementing other arms control measures.

General Secretary Khrushchev, while criticizing the proposition for limiting the moratorium to one year and for its two accompanying conditions, accepted the October 31 date for beginning negotiations. In the intervening two months the USSR, US, and UK each undertook an extensive series of tests.

The diplomatic negotiations, again conducted in Geneva, began on October 31, 1958, but three weeks were consumed in a dispute over the formal conference agenda, until the delegates gave up and proceeded without such an agenda. Even the title of the talks, "The Conference on the Discontinuance of Nuclear Weapons Tests," was in dispute, with the USSR equating "discontinuance" with "cessation," and the West interpreting it to mean "suspension."

After several weeks the USSR introduced a draft treaty of five short articles. It bound the three powers not to undertake nuclear tests in any medium, and to discourage the commencement of nuclear testing by all other states in the world. Compliance would be verified through the detection network recommended by the Conference of Experts.

The Western delegates rejected the draft on grounds that no mention was made of other disarmament measures, no control organization was specified, and no provision was made for sanctions against violators. Furthermore, the American and British negotiators were still less than fully confident in the capabilities of the Geneva System of the Conference of Experts. After several more weeks the USSR agreed to the inclusion of a control commission provision, although the exact composition of such a body remained in dispute when the negotiations recessed on December 19.

Upon the resumption of talks on January 5, 1959, the chief US delegate, James Wadsworth, held an informal meeting with his Soviet counterpart, Semyen Tsarapkin. He informed him that in the course of underground tests conducted in Nevada in October 1958, data were obtained which indicated that the detection of tests would prove more difficult than previously believed. Based on these tests, known as the HARDTACK II series, it appeared that the number of earthquakes equivalent to a given explosive yield would be far greater than earlier estimates, perhaps by a factor of ten to fifteen. In addition, better data on background noise indicated that the first motion would be more difficult to determine accurately. These findings meant, according to Wadsworth, that the 90% confidence identification threshold proposed in the Geneva System would be more on the order of 20 kt than 5 kt, and that a far larger number of control stations would be

required to maintain the 5 kt threshold.

In the next few weeks other ideas were put forward which further dampened hopes of concluding a treaty. Carl Romney, a seismologist for the US Air Force, stated in congressional testimony that, based on HARDTACK data, the number of earthquakes indistinguishable from 5 kt explosions would not be 20 to 100 per year, as estimated at the Conference of Experts, but probably 700 to 3000, of which 100 to 500 would be in the USSR and China [10]. Romney's 1959 congressional testimony on monitoring capability marks one of those points in the early CTBT debates where the technical assessments offered to policymakers turned out in retrospect to be way off track. As became apparent by the early 1970s, the number of earthquakes indistinguishable from 5 kt explosions in the USSR and China were few in number – in fact, approximately zero per year, not 100 to 500. Romney's summary statement drew upon answers to several separate technical questions, such as:

- How many earthquakes occur each year at different magnitude levels?
- What would be the magnitude levels down to which a hypothesized global network of seismographic stations could achieve reliable detection and identification?
- What was the relationship between the yield of an underground nuclear explosion, and its seismic magnitude?

In each of these areas, Romney's 1959 testimony offered numbers quite different from what was discovered and accepted in later years. Thus, he concluded there were about 20,000 earthquakes worldwide, each year, whose seismic magnitude was 4 or greater. Today, we would give about 7500 for this number [11]. He concluded that what today would be considered a dense global network (the Penney proposal) could supply reliable detection only down to the magnitude range of about 4 to 4.5. Today, we would give about 3 to 3.5 for this magnitude range, even with a less dense network (paper CD/1254, of the Conference of Disarmament). In this early period he was apparently using the relation [12, Fig. 15]:

$$m = 3.9 + (2/3) \log Y \text{ (in kt),}$$

which gives a yield (Y) of about 19 kt corresponding to a magnitude (m) of 4.75, whereas this magnitude was later found to correspond to about 10 kt for a typical underground explosion in wet tuff at the Nevada Test Site (see [13], which in 1981 reported the relation for this test site as  $m = 3.92 + 0.81 \log Y$ ), and to only 2.5 kt for a typical explosion at what became the USSR's main test site (see [14], which in 1992 reported  $m = 4.45 + 0.75 \log Y$  for the Balapan area of the Semipalatinsk Test Site).

Going back to our historical review, several points should be noted in connection with the HARDTACK data. Out of eight underground explosions included in the series, only two, BLANCA (19 kt) and LOGAN (5 kt) produced seismic data sufficient for evaluation of system capabilities. US seismologists also concluded that the previous magnitude estimate for the 1957 RAINIER test was too high because the seven best

stations near the event had given what now appeared to be anomalously large results. Thus, RAINIER, known to be 1.7 kt in yield, should have been estimated at magnitude 4.06 ( $\pm 0.4$ ) and not 4.25 ( $\pm 0.4$ ) as previously believed. In addition, difficulty in assessing signs of first motion in the HARDTACK tests indicated that the ratio of first motion to background noise amplitude must be at least 3 to 1, rather than 2 to 1 as estimated earlier [15].

Just as the significance of HARDTACK was being debated, a theory even more discouraging to test ban advocates emerged. Albert Latter, of the RAND Corporation, presented preliminary findings on the principle of cavity decoupling – the explosion of a bomb in an underground cavity large enough that the surrounding rock would not deform plastically (permanently) in any direction, but remain elastic. Under such conditions, according to Latter, the seismic signal could be reduced by a factor of as much as 300, thereby rendering impossible the detection of all but the very largest tests [16].

The controversy surrounding HARDTACK and cavity decoupling led James Killian, the President's Special Assistant for Science and Technology, to appoint a Panel on Seismic Improvement, chaired by Lloyd Berkner. The Berkner Panel, in a report on March 16, 1959, concluded that, given current technology, an increase in the number of seismometers at each array station from 10 to 100 would make detection of the first motion much easier. They also recommended the establishment of supplementary unmanned stations, 170 km apart in seismic areas, which could identify 98% of one kiloton events. Future improvements were discussed, including the use of seismometers in deep boreholes, an extensive chemical explosive testing program, and computer-aided reconstruction of wave forms. Regarding cavity decoupling, the Berkner Panel reached a less pessimistic conclusion than Dr. Latter, and stated that apparent yield could probably be decreased ten-fold, depending on the surrounding rock type [17].

The Berkner Panel also recommended major funding for research in basic seismology, with particular attention to improvements in data acquisition. Thus, a key recommendation was that "100-200 of the existing stations in the world be equipped with modern instruments as soon as possible." Annual budgets of about \$18 million were outlined for such improvements, plus another \$12 million for each of two years to carry out chemical and special nuclear explosions underground for monitoring research [18]. Since, prior to 1960, the field of pure seismology in the United States had received national support at the level of only about \$0.7 million annually [19, p. 80], and since funding at the level recommended by the Berkner Panel (for seismology – not for nuclear explosions) was actually appropriated and spent [19, pp. 27-37], verification research has had an enormous impact on seismology – and on geophysics in general.

In 1959, many feared these new US reports would cause the Geneva talks to collapse. The Soviets refused to consider the HARDTACK data on procedural grounds, namely that the seismometers used were not identical to those specified in the Conference of Experts. The US responded that in a study of small chemical explosions the

HARDTACK seismometers had outperformed the Geneva-specified equipment. Decoupling theory was greeted with outrage on the part of the Soviets, who asked why a nation serious about concluding an agreement should devote time and money to devising means of circumventing it.

Torn between his own desire for a test ban and this new evidence, Eisenhower proposed on April 13, 1959 a treaty to ban only “policeable tests,” defined as those in the atmosphere up to 50 km and in the ocean. Khrushchev vehemently rejected the suggestion, claiming that the US would continue nuclear testing underground and in space and that, in any event, all tests were “policeable.”

Despite their reservations, the Soviets agreed to the establishment of Technical Working Group I, in the summer of 1959, to study high altitude and space monitoring. In the meantime, debate continued on the number of on-site inspections necessary to ensure treaty compliance. The USSR, although specifying no exact figure at this time, insisted the number should be determined by political considerations, while the West held to inspections automatically triggered by technical criteria. Due to Khrushchev’s upcoming visit to the US, as well as general elections in the UK, the negotiations recessed in 1959 from August 26 to October 27.

Soon after the Conference resumed Tsarapkin proposed that a second working group be convened to examine the controversial new seismic data introduced by the US. The first presentation made in Technical Working Group II was by Carl Romney, and concerned the HARDTACK II data. While including more details, he basically concurred with the position presented in January, that the magnitude of RAINIER had been overestimated and therefore many more earthquakes of equivalent size existed. The Soviet response was couched in legalistic terms, and again charged that no conclusions could be drawn regarding the control system since the seismographs used in HARDTACK II were not identical to those recommended by the Conference of Experts.

Although this charge in retrospect may seem irrelevant, other Soviet criticism was more substantive. To begin with, less than thirty of the stations which recorded BLANCA and LOGAN were sufficiently calibrated to measure magnitude. Data scatter was highly controversial as well. The Americans insisted that all points should be used in computing an average, while the Soviets argued that points in the so-called “shadow zone,” of intermediate distance, should be excluded. The background to this argument concerns a property of the Earth of great importance for seismic monitoring, associated with a zone of lower seismic velocities hundreds of kilometers deep in the upper mantle, that has the effect of defocussing seismic waves received in the distance range about 1000 to 2000 km from an earthquake or an explosion. The property is illustrated in Figure 1, which shows schematically the way in which the amplitude of seismic waves at first decreases, and later increases, as the waves propagate to increasing distances from a shallow earthquake or explosion. The region of low amplitude, or lack of

observations, is the shadow zone. At lesser distances (the “First Zone”), or greater distances (the “Third Zone”), amplitudes are larger and are thus more likely to be routinely observed above the noise. The inclusion of low amplitude signals within the shadow zone tended to lower the seismic magnitude assigned to BLANCA and LOGAN.

*Figure 1.* The effect of variation in seismic wave speed with depth in the Earth (see bottom right), upon the path of propagation of the fastest seismic body waves (lower section, showing ray paths of propagation in the crust and upper mantle). Amplitudes are relatively strong, out to distance ranges of around 1000 km (the First Zone); are weak or absent in the range around 1000-2000 km (the Shadow Zone); and become strong again beyond 2000 km (the Third Zone). For the decades following the 1963 signing of the LTBT, when nuclear testing was carried out underground but no in-country verification was permitted, monitoring was conducted by National Technical Means using seismic signals acquired typically in the Third Zone. Nomenclature changed, such signals came to be called “teleseismic waves,” and they have been intensively studied especially for purposes of yield estimation in the context of the Threshold Test Ban Treaty. With renewed attention to CTBT verification, and the use of in-country stations, it has become more important to return to the study of the strongest signals, acquired in the First Zone. Here too the name changed, and such signals are more commonly now referred to as “regional waves.” From [12].

The most dramatic moment in TWG II occurred when Hans Bethe and Albert Latter officially presented the theory of cavity decoupling. Bethe later commented [20] that “the Russians seemed stunned by the theory of the big hole,” a concept which “implied that we considered the Russians capable of cheating on ... a massive scale.” After several meetings the Soviets admitted that Latter’s work was theoretically valid but argued there was no proof it could be made to work in practice. The British delegation presented the results of several small tests of TNT in cavities, conducted during the summer, which seemed to support the theory. The participants were unable to devise

any recommendations at this time for foiling the decoupling strategy.

The final item covered in TWG II was the formulation of criteria to initiate on-site inspections. The Soviets proposed that if the epicenter of an event were located in an area of dense population, or if its focal depth were beyond current drilling feasibility, it would be ruled an earthquake. Only events exhibiting positive explosion characteristics should be considered for on-site inspections. The Western delegates, stating that no such positive criteria existed, introduced counterproposals that, according to the Soviet side, would have made the majority of recorded events open to suspicion and eligible for on-site inspection. An American proposal listed several potential methods which might, after sufficient research, evolve into criteria for establishing an event as an earthquake.

In the end, technical differences were too deep to allow TWG II to issue a joint statement. The participants agreed instead on a report detailing the proceedings, with four attached annexes. The first listed some generally agreed recommendations on improvements, while the other three consisted of the differing views of the Soviet, British, and American delegations.

On February 11, 1960, the US presented a new position at Geneva. It proposed a phased treaty which would immediately prohibit tests in the atmosphere, underwater, and underground down to the lowest adequately controlled threshold. This threshold was proposed to be 4.75 on a unified magnitude scale. According to data from Nevada (see above), the authors of the proposal felt that this limit would correspond to about 19 - 20 kilotons, fully coupled (i.e., tamped in the surrounding medium so that seismic waves would be excited efficiently). In addition, the proposal accepted the idea of a numerical quota of on-site inspections, probably 30% of unidentified events, which the authors felt would be sufficient to deter cheating. On average this was believed to correspond to twenty inspections in the USSR per year, and an equal number in the US and UK together [4, p. 16]. The proposal concluded with a suggestion that the three nations undertake joint seismic research to reduce the threshold further.

The problem of proliferation, often cited as a strong reason for concluding a test ban, re-surfaced two days after this proposal. On February 13, 1960, France exploded a nuclear device, of 60 to 70 kiloton yield, in Algeria, thereby ignoring recommendations of the UN General Assembly.

The USSR replied to the new US position on February 16. Tsarapkin at first rejected the phased treaty concept since the USSR favored a comprehensive ban. He proposed several temporary identification criteria, which might apply for an initial period of two to three years. An event could be inspected if data from several surrounding stations localized it to an area of 200 square kilometers, but would be ineligible if its depth of focus were found to be more than 60 kilometers, its epicenter were found to be oceanic with no accompanying hydroacoustic waves, or if it were established within 48 hours as the foreshock or aftershock of an earthquake. However, on March 19,

Tsarapkin issued a more detailed proposal, in which the USSR announced a willingness to conclude a treaty “on the cessation of all nuclear weapon tests in the atmosphere, in the oceans, and in outer space, and of all underground tests which produce seismic oscillations of magnitude 4.75 conventional units or above” [21]. He agreed to a joint research program. While accepting most elements of the US proposal, the Soviet plan contained no on-site inspection quota, and insisted on a moratorium on tests below magnitude 4.75.

In light of later seismic data and later negotiation of the Threshold Test Ban Treaty, the acceptance by the USSR of a magnitude threshold is noteworthy. At the time of the acceptance, magnitude 4.75 was regarded by US experts as corresponding to about 20 kt for an underground explosion in Nevada under conditions of the RAINIER explosion, though later the yield equivalence was found to be about 10 kt (discussion of Romney testimony; see above). And, as also noted above, magnitude 4.75 turned out to correspond to only about 2.5 kt at the site the Soviets later developed for underground testing in East Kazakhstan. In retrospect, a threshold test ban based on magnitude would therefore have restricted yields on the East Kazakhstan Test Site significantly more than yields at the Nevada Test Site. Some early hints at the conclusion that magnitude-yield relations might vary between test sites were in fact available in 1960, for although the USSR had not yet conducted an underground nuclear explosion, the Soviets had experience with large underground chemical explosions. For example, on March 3, 1960, 660 tons of TNT were placed in an underground chamber and fired as a single charge at depth in Kirgizia (present-day Kyrgyzstan), resulting in seismic waves so strong that they were reported to the US as the equivalent of 5 kt fired “under RAINIER conditions.” Today we would expect such a chemical explosion in Kirgizia, and a 5 kt RAINIER-type nuclear explosion in Nevada, both to have magnitude around 4.5. But in 1960 Albert Latter stated that “I personally do not accept the Russian statement because they have not given any confirmatory details” [22].

Several US research programs relating to underground verification were carried out in 1960. Project COWBOY, conducted in March, consisted of a series of chemical explosions set off in cavities, and generally supported the Latter brothers’ decoupling hypothesis, although indicating a decoupling factor lower than 300. Accompanying studies concluded that the technology to construct such cavities, through solution mining in salt domes, already existed. Project VELA, the research effort in seismology recommended by the Berkner Panel, began in 1959 under the Advanced Research Projects Agency of the US Department of Defense, and by 1960 had developed detailed plans for evaluating detection and identification capabilities using small chemical and nuclear explosions. The USSR protested that adequate safeguards must exist to prevent these small nuclear shots from being used for weapon development purposes – and indeed the US DoD had plans to couple such supposedly seismic experiments with installations to

study weapons effects [23]. But although the chemical explosions under VELA were carried out, the nuclear explosions were postponed indefinitely due to a lack of agreement on safeguards [4, p. 265].

Project VELA also included a plan by the US Coast and Geodetic Survey to build a World-Wide Standard Seismograph Network (WWSSN), following the recommendation of the Berkner Panel. The network would shortly grow to include about 125 stations, most outside the US, each recording in a standard analog format on photographic paper. The stations recorded continuously, and this network collected earthquake data around the world as well as nuclear explosion data. The WWSSN has had a profound influence on the growth and achievements of the science of seismology, providing important support and insight into the theory of global tectonics which revolutionized the Earth Sciences in the 1960s. In addition to the WWSSN, a decision was made to build seven small arrays in the US that conformed to recommendations made in the Conference of Experts, and designed explicitly to detect Soviet tests. The first of these, at Fort Sill, Oklahoma, appeared capable of detecting events down to magnitude 4, equivalent to a one kiloton shot, at distances “over 2000 miles,” (i.e. in the Third Zone) although it appeared identification would not be reliable until the event approached the equivalent of 5 kilotons [24].

Meanwhile, the Geneva Conference turned to the question of how many control posts would be needed and where they should be situated. On May 12, 1960, the US proposed that in the initial phase of a three phase process, 21 posts should be constructed in the USSR, 11 in the US, 1 in the UK, 2 on ships, and 12 on islands in the Northern Hemisphere, for a total of 47. (In Washington two days earlier, Secretary of State Christian Herter had been horrified to learn of a study sponsored by the Department of Defense which estimated it would take \$1-5 billion to install the 21 control posts in the USSR. The plan turned out to include the building of large airfields by the US in the USSR, and hiring icebreakers to take in supplies. The estimate was found to be inflated and the study declared invalid a few weeks later [23, pp. 323 & 348].) The USSR complained in Geneva that this scheme did not provide any posts in the Southern Hemisphere, where the Western powers often tested. The Soviets instead advocated 15 posts in the USSR, 11 in the US, 1 in the UK, 7 in Australia, 20 on islands belonging to the UK and US, 2 in Canada and/or Mexico, 2 in Africa, and 10 on ships, for a total of 68 posts. In addition, the Soviets insisted that no on-site inspections take place during the first phase of installation, which would probably take four years. The Western powers argued for dividing this phase into two two-year periods, and beginning inspections at the end of the first period. At the time of the US election in November 1960, when John Kennedy defeated Vice-President Richard Nixon, these differences remained unresolved.

Figure 2 here shows the locations of stations (arrays) in a global network of the type

discussed in Geneva in 1960; and, for comparison, one of the designs proposed in 1994 for a test of global network operation (known as GSETT-3: see elsewhere in this volume) that is to begin in 1995. Both of these networks are hypothetical at this time of writing, though practical experience with the later design (or something like it) will be available in 1995. The sparse 1994 design is predicted to have a significantly better detection capability, than was associated in 1960 with the global Geneva system.

*Figure 2.* Hypothetical global seismographic networks. Upper: a design proposed in 1960 [19, p. 58] and based upon the Penney proposal of 1958. Each continental post was to be an array of about ten stations. Lower: a design proposed in 1994 as a mix of about 20 arrays and 30 single stations. Networks of this latter type are claimed to provide detections, adequate for location, down to about

magnitude 3 - 3.5 in much of Eurasia and North America.

Upon assuming office, President Kennedy undertook a thorough reorganization of the US arms control apparatus. A new unit of the State Department, the Disarmament Administration, was created. Arthur Dean was named to replace James Wadsworth as chief representative to the Geneva Conference, and AEC Chairman John McCone was replaced by Glenn Seaborg, who was more inclined to favor to a test ban.

The new US position, presented when the Geneva Conference resumed on March 21, 1961, contained a few minor concessions. The US would now seek legislation permitting the Soviets to examine the internal mechanisms of nuclear devices employed in US seismic research and peaceful explosions programs [25, p. 56]. The US proposal continued to insist on a quota of twenty inspections per year in the USSR, as opposed to the Soviet proposal for three, but was willing to assign quotas of twenty inspections to the US and UK as well. Very few modifications were made in the technical issues, however, and the US still envisioned a threshold set at seismic magnitude 4.75.

The Soviet reply was pointed and negative. Tsarapkin denounced the testing of weapons by France as a serious obstacle to progress, and accused the US of dragging negotiations out long enough to shift research work for NATO into French hands.

On August 28, 1961, virtually as a desperation measure, Ambassador Dean offered to eliminate the 4.75 seismic magnitude threshold if the USSR would agree to an increase in the number of control posts or on-site inspections. As expected, the USSR rejected this proposal. Three days later the Soviet Union ended its moratorium and conducted the first test of what would be its most extensive series ever – a series which had obviously been in preparation for some time. The accompanying statement minimized the importance of a test ban alone, and used the French tests and current Berlin Crisis as pretexts for resuming testing. The US Atomic Energy Commission reported that atmospheric nuclear explosions in the kiloton range took place at the Semipalatinsk Test Site, in East Kazakhstan, on September 1, 4, 5, 13, 17, and October 12, 1961; and east of Stalingrad on September 6. On October 11, 1961, the Soviet Union's first underground nuclear explosion took place, also on the Semipalatinsk Test Site. It had magnitude about 4.8 [26]. This explosion was detected at six stations of the USCGS's new worldwide network and at one Swedish station, and was apparently identified as underground and nuclear [27], although the event was not widely listed as the USSR's first underground nuclear explosion until the 1980s.

The US was ill-prepared to resume testing – Los Alamos and Livermore had not even been allowed to buy cable since this might have signalled an intention to break out of the moratorium [28] – but began with a small underground test on September 15 while still refraining from atmospheric tests. The Soviet Union, despite a UN resolution calling on it to refrain from a proposed atmospheric test of 50 megatons or more, exploded the largest atomic device ever tested on October 30. Its yield was estimated at

58 megatons, but Hans Bethe speculated that if its fusion material had been encased in uranium rather than lead the yield could have been in excess of 100 megatons [29].

Following a recess in the Geneva Conference during October and November, the Soviet Union introduced a proposal for the immediate conclusion of a treaty banning space, atmospheric, and underwater tests, and a moratorium on underground tests pending an agreement on a control system. The US and UK rejected any proposal omitting a specific control system and, in the absence of any further progress, the Conference finally ended on January 29, 1962, without the release of any joint communiqué.

The collapse of the Geneva Conference coincided with the creation of a panel, headed by Hans Bethe, to evaluate the most recent Soviet test series. This panel stated that the USSR had made sizeable gains in reducing the weight to yield ratio of its weapons, in increasing overall yield, and in reducing the size of the necessary fission trigger. The panel also concluded that much of the preliminary research for this series was conducted during the three year moratorium on nuclear tests.

In the meantime, new seismic data became available regarding explosions in various media. On December 10, 1961, Project GNOME, the explosion of a 3 kt nuclear device in a salt dome in New Mexico, was conducted. Based on results obtained in Project COWBOY, it had been believed that a fully tamped shot in salt would produce a signal smaller, by a factor of perhaps two and one half, than a tamped explosion in tuff, the rock type in which all previous US underground tests had been conducted. Contrary to these expectations the signals from GNOME were significantly larger than those of LOGAN, a 5 kt shot tamped in tuff at the Nevada Test Site in 1958 [4, pp. 351-352]. The GNOME shot was detected as far away as Japan and Sweden. This was the first clear indication to the US, that the relation between magnitude and yield could vary significantly from one region to another. The reasons have to do with the differences in rock type in the immediate vicinity of the shot point (which affect the efficiency with which explosion energy is coupled into the energy of seismic waves), and the differences in propagation characteristics of seismic body waves in different regions (which affect the way in which body waves are attenuated, as they travel from the seismic source to stations at which the signal strength is recorded).

However, although the fact of the stronger-than-expected GNOME signals was encouraging to those seeking effective ways to monitor underground nuclear explosions, other results from this shot were less encouraging. The discovery was made that seismic wave velocities through the Earth's crust were not uniform from one region to another, making more difficult the analysis of signals to obtain a source location. Had the USSR proposed position on inspection criteria been in force, a 200 square km area around the estimated epicenter, the GNOME shot would have occurred outside the area eligible for inspection. Furthermore, the depth of the GNOME event was not estimated near 350 meters, the actual depth of detonation, but rather at about 130 km, which would have identified it as an earthquake. In general, these uncertainties led many to

lose confidence in the capability of seismological methods to verify a nuclear test ban effectively. By making appropriate corrections for the non-uniformity of the Earth's crust, event location could still be done accurately. But what would be the confidence in the corrections, for an event in an area where the corrections had never before been derived, and the "ground truth" data to do so were unavailable?

On February 2, 1962, the US Atomic Energy Commission announced that earlier that day the USSR had apparently conducted an underground nuclear test. The test, widely reported to be the first underground Soviet nuclear explosion, was carried out in a generally aseismic area in Soviet Central Asia (East Kazakhstan), and had a yield estimated at 40 to 50 kilotons. The rapid detection and rapid identification of this test were applauded by proponents of the test ban [4, p. 353], few (if any) of whom knew that there were data, available in the West, to indicate that a previous underground nuclear test had taken place in the USSR (see above).

The section of Project VELA concerned with cavity decoupling, Project DRIBBLE, was planned to consist of six explosions, both tamped and decoupled. Due to lack of funds this project was temporarily suspended after exploratory drilling and engineering work. Upon its resumption in September 1962 a cavity for a 100 ton shot was planned which would require a year of work and cost \$3.2 million. By 1965 more than this had been spent and construction of the cavity had not yet commenced – an indication, presumably, of the difficulty in executing a decoupled explosion, even without the additional problems of keeping the shot secret [30, p. 312].

Within two months of the end of the Geneva Conference, international pressure, especially an emotional appeal from Prime Minister Harold Macmillan, led the parties back to the negotiating table. The forum was now multi-lateral, and was called the Eighteen Nation Disarmament Committee, consisting of five NATO states, five from the Warsaw Pact, and eight non-aligned states. The ENDC, a forerunner of today's Conference on Disarmament, convened on March 11, 1962, and began with the Soviet Union tabling a draft treaty on General and Complete Disarmament. The Soviets attributed the failure of previous negotiations solely to US intransigence, and went on to claim that National Technical Means – surveillance that a country could unilaterally achieve without access to the territory of the country being monitored – would be sufficient to detect underground as well as atmospheric tests.

The US response consisted of a test ban proposal incorporating four modifications to previous Western demands. First, to prevent surprise abrogation of the treaty, heads of state would make periodic declarations that no test preparations were underway, and declared test sites could be inspected by the other party a certain number of times per year. Second, the inspection process and the establishment of control posts were to be inaugurated sooner than the two years previously discussed. Third, the 4.75 magnitude threshold was to be eliminated due to the difficulty of determination, making the treaty

comprehensive. Fourth, on-site inspections would mainly be confined to a normally aseismic area in Siberia, with only a few in the heart of the USSR. Although the last two provisions were considered by the US to be major concessions on its part, the USSR rejected this proposal, arguing again that National Technical Means must suffice for any treaty.

Despite appeals from the non-aligned states as well as several allies, the US resumed atmospheric testing on April 26, 1962. This series, which included a few proof tests of existing stockpiles as well as new weapon development, totalled about twenty megatons yield. Totalling the activities of the US, USSR, UK, and France, more nuclear weapons were tested in 1962 than in any other year, and more total megatonnage detonated, from September 1961 to December 1962, than in any other period of comparable duration.

While the ENDC was stalemated, several developments in the US increased support for an atmospheric test ban. During the spring of 1962, following large test series by the US and the USSR, the level of fallout-induced radioactivity was found to have increased significantly worldwide. Several scientists proposed that the concentration of iodine-131 in the atmosphere had reached dangerous limits, and that protective measures might be necessary for some foodstuffs, especially milk, if tests continued at the same rate. At the same time, US nuclear strategy was officially stated to be changing from one of massive retaliation to a doctrine of targeting Soviet weapons systems. Very large warheads thus became less desirable, reducing the need for atmospheric tests.

During this period Project VELA began to produce a more informed understanding of monitoring capability, and several useful and specific results. An underground French test, conducted in Algeria on May 1, 1962, was detected by several of the new Coast and Geodetic Survey stations, and estimated at 30 to 50 kt. This test, like the Soviet tests in East Kazakhstan, indicated the feasibility of teleseismic detection (i.e., data that had been acquired in what earlier was called the Third Zone, beyond the shadow zone, see Figure 1, and thus available by National Technical Means). Next, the discovery was made that previous estimates of the annual number of shallow earthquakes in the USSR were too large. These estimates, based on extrapolations of earthquake records from 1932 and 1936, had indicated 100 shallow earthquakes above magnitude 4.75 (then thought to be equivalent to 19 or 20 kt in tuff), and 600 above magnitude 4.0 (then thought to be 2 kt in tuff). Using more recent and better data the figures were revised to about 40 shallow earthquakes above magnitude 4.75, and 170 above magnitude 4.0 [31].

The placement of seismometers in deep boreholes was soon found to increase their sensitivity (specifically, their signal to noise ratio) by a factor of five or ten. In addition, special filtering of data from surface arrays of many seismometers was seen to improve capabilities considerably. Finally, research showed that seismometers positioned on the ocean floor could provide useful monitoring data.

Some other developments, however, indicated new difficulties in detection. Seismic

signals measured in different directions from an explosion were found to be of significantly different strengths. American scientists also discovered that a test carried out in loosely compacted alluvium would produce a signal only one-seventh as large as a test in tuff (one-fourteenth as large as one in granite). However, an underground test in alluvium would most likely cause a cavity visible on the surface.

Overall, the Project VELA results were encouraging, and the US felt confident enough to introduce two new draft treaties in Geneva on August 27, 1962. In the eyes of the world, the US position was enhanced by the commencement of a new Soviet test series on August 5, which included a 30 megaton shot. The first draft for a Comprehensive Test Ban Treaty envisioned a fifteen member International Scientific Commission (four Western, four Eastern, seven non-aligned) to establish standards for the calibration and operation of all elements of the verification system. This system would consist of nationally owned and manned stations as well as several new facilities financed and staffed by the Commission, to be constructed at sites listed in an annex to the treaty. Equal quotas of on-site inspections would be assigned to the territory of the USSR, UK, and US. Any event not positively identified as an earthquake, by first motion or depth, would be eligible. No number was specified at this time for inspections, but the UK and US delegates stated it would be less than the 12 to 20 previously proposed. Data from Project VELA now indicated that only about 10 to 15 unidentified events of magnitude  $\geq 4.75$  would occur in the USSR each year [32, p. 15].

The second Western draft was a far briefer proposal for a Partial Test Ban. This treaty prohibited tests in or above the atmosphere, in the seas, and in any other environment if the explosion caused radioactive debris to escape outside the territorial limits of the testing state. The last provision was intended to prevent a nation from putting a small amount of earth over a surface shot and styling it an underground test. The draft did not mention the creation of a control system or international organization, nor did it call for any moratorium on underground tests. Ambassador Dean, when presenting the Partial Ban draft, declared it could and should be accepted immediately as a means of limiting the arms race and stopping radioactive pollution.

Detection capabilities under both proposed treaties were evaluated by a panel headed by Assistant Secretary of Defense Paul Nitze. He estimated that the system envisioned in the comprehensive draft could detect underground shots down to about 10 to 20 kt in alluvium, 1.5 to 3 kt in tuff. Nitze stated that this threshold would still allow the USSR to study most important technical principles of nuclear weapons development, including those relating to neutron weapons. His panel concluded that detection capability for atmospheric and underwater tests was adequate, but that tests conducted in inland waters or outer space would be difficult to detect [33].

The USSR rejected both treaties, the first because it still allowed for on-site inspections which the West could use for espionage purposes, and the second because it

permitted underground tests. The eight neutral states in the ENDC sought to placate the Soviets by proposing that the entire International Scientific Commission decide which suspicious events should be inspected, rather than the opposing nuclear power. The Western powers rejected this suggestion and the ENDC recessed on September 7.

In October 1962, the crisis over Soviet missiles in Cuba convinced both superpowers of the need for rapprochement. Having confronted the very real possibility of nuclear war, the USSR and the US were more willing to moderate negotiating positions. The Soviet Union indicated it would be willing to consider the use of sealed automatic recording stations, nicknamed “black boxes,” for in-country verification, based on a suggestion by three American and three Soviet scientists at the Tenth Pugwash Conference in London. Ambassador Arthur Dean revealed to the Soviets that the US might now accept 8 to 10 on-site inspections per year, and 8 to 10 nationally manned control posts, in the territory of the USSR.

When the ENDC reconvened in November the USSR proposed the use of automatic seismic recording stations to eliminate the need for internationally supervised, nationally manned stations as well as on-site inspections. Several weeks later they suggested three possible sites for these “black boxes,” and announced they would be willing to have international personnel participate in the installation of these devices on Soviet territory. Although the Western powers rejected the idea of eliminating on-site inspections, they proposed that a group of experts be convened to discuss the “black boxes.”

Following this proposal Kennedy and Khrushchev exchanged letters discussing the acceptable number of on-site inspections in the USSR. Kennedy advocated 8 to 10, while Khrushchev demanded 2 to 3. Private talks were then conducted in the US between William Foster, director of the Arms Control and Disarmament Agency (created in September 1961 in President Kennedy’s new Administration), and Soviet representatives. Potential “black box” sites in both countries were proposed and accepted or rejected, and seismic noise level data for the sites exchanged. The US felt its requirements might be satisfied by as few as seven such stations, but when the talks ended on January 31, 1963, the Soviets were willing only to consent to three [25, p. 184]. The same day the talks ended, Edward Teller presented a paper to a group of influential Republican Congressmen, charging that acceptance of current Soviet proposals would be equivalent to accepting an unpoliced moratorium.

Due to the unresolved inspection issue, opposition to a test ban developed in the US – or at least to a CTBT. In March 1963 hearings were conducted by the Joint Committee on Atomic Energy to discuss the technical aspects of verification. Carl Romney testified that the seismic system under consideration by the US would be able to detect most tests down to 1 kt in granite, 2 to 6 kt in tuff, and down to 20 kt in alluvium. However, he stated that decoupling could attenuate seismic signals by a factor of 200 [30, p. 104]. Much of the criticism of the Kennedy administration’s

determination to conclude a treaty came from the Republican Conference Committee on Nuclear Testing, chaired by Rep. Craig Hosmer, who had commanded the first occupation troops in Hiroshima in 1945. It appeared that, even if an inspection number acceptable to the Soviets were found, not enough support for a comprehensive ban existed in Congress to provide the Senatorial advice and consent required to ratify a Comprehensive Test Ban Treaty. Apparently Kennedy hoped to generate as much support for the treaty in the Senate as possible, and wanted to receive more than the minimum two-thirds vote needed for ratification. According to Glenn Seaborg, then the Chairman of the AEC, “Kennedy felt that the treaty needed to be launched on a strongly positive note to serve its purpose as a first step to a better world order” [25, p. 258].

On July 2, Khrushchev announced that the USSR would be willing to accept a treaty banning tests in the atmosphere, in space, and underwater. For the first time, he did not insist that an underground moratorium accompany the treaty. The following day US officials replied that the Administration would also accept such a partial ban. President Kennedy dispatched the veteran diplomat W. Averell Harriman to Moscow with broad instructions to attempt to conclude a comprehensive ban but to settle for a partial ban if necessary [25, p. 229].

After all the preliminaries, the final steps to conclusion of what has become known as the Limited Test Ban Treaty were anticlimactic. The negotiations began on July 15, 1963, with the US making a final attempt to negotiate a comprehensive ban. An effort was made to arrange meetings between Frank Press, the only seismologist in the US delegation, and Soviet seismologists, but it was claimed that these were all away from Moscow or otherwise unavailable [4, p. 455]. A few days later this attempt was abandoned, and a draft, based on the Western proposal for a partial ban issued the previous year, was put forward.

A treaty, virtually identical to this draft, was composed, and signed on July 25 by Foreign Minister Andrei Gromyko, Ambassador Harriman, and Science Minister Lord Hailsham. In five short articles, it prohibited testing at sea, in the atmosphere, in space, and in other environments if such tests caused radioactive debris to be present outside the testing state’s territory. The treaty was to be of unlimited duration, and would be open to all states for signature. Signatory states would have the right to withdraw, if extraordinary events jeopardized their supreme interests, with three months notice. Remarkably, no mention whatever was made of verification systems or international control, it being assumed that National Technical Means would suffice.

On August 5, the treaty was signed by Foreign Minister Gromyko (for the second time), Secretary of State Rusk, and Foreign Secretary Lord Home. Three days later it was submitted to the US Senate for approval. Kennedy felt that the endorsement of the Joint Chiefs of Staff, as well as a majority of the scientific community, was essential to secure the consent of the Senate. General Maxwell Taylor, Chairman of the Joint Chiefs of Staff, testified that four crucial safeguards were necessary for the military to

recommend ratification of the treaty:

- An extensive underground test program must continue in order to improve the US arsenal.
- Modern nuclear laboratory facilities and research programs must be maintained.
- The resources and facilities to resume atmospheric tests promptly must be maintained, in the event of Soviet non-compliance with the treaty.
- US capability to monitor the treaty and detect violations must be improved.

President Kennedy, in private conversations with the Chiefs, supported these safeguards. In the view of Glenn Seaborg: “While this support may have obtained the favorable testimony of the Joint Chiefs, it was at a very heavy price for the cause of disarmament” [25, p. 271]. At the end of these hearings, which lasted three weeks, the Joint Chiefs of Staff gave their formal, if unenthusiastic, approval to the limited ban.

As expected, Edward Teller opposed the treaty, arguing that the US needed to test in the atmosphere to learn more about weapons effects. Addressing the Senate, he stated that if they consented to ratification, “You will have given away the future safety of this country. You will have increased the chances of war, and, therefore, no matter what the embarrassment may be in rejecting the treaty, I earnestly urge you to do so and not to ratify the treaty which is before you” [34, p. 428]. John Foster, Director of the Livermore Laboratories, considered the treaty “disadvantageous from purely technical-military considerations,” and urged rejection. The Director of Los Alamos, Norris Bradbury, supported the treaty, but only on condition that the US government devote itself to a vigorous underground test program. Hans Bethe and other test ban proponents, while regretting failure to negotiate a comprehensive ban, applauded the treaty as a useful first step [34, pp. 583 & 616].

Ratification of the Moscow Treaty, formally known as the Treaty Banning Nuclear Weapon Tests in the Atmosphere, Outer Space, and Underwater, received the consent of the US Senate, by a vote of 80 to 19, on September 29. The Presidium of the Supreme Soviet voted unanimously to ratify the treaty on September 25. On October 7 President Kennedy signed the Moscow Treaty, which entered into effect on October 11. That same day, in Oslo, the Nobel Peace Prize was awarded to Linus Pauling, one of the very earliest test ban advocates.

The Moscow Treaty, commonly referred to as the Limited Test Ban Treaty, promptly found wide support. By the end of 1963, 113 nations had added their signatures to those of the USSR, US, and UK, and by late 1994 over 145 states had become signatories. The People’s Republic of China, which was to test its first atomic bomb in 1964, and France, which felt that its independent arsenal, the *force de frappe*, still needed perfection through atmospheric tests, were the most prominent states refusing to sign. However, although the LTBT was successful as an environmental measure in that radioactive fallout was greatly reduced (even France and China eventually stopped

atmospheric testing – the last such test being conducted by China in 1980), the treaty had little impact on nuclear weapons development in view of the vigorous programs of underground testing that continued for decades. The work to obtain agreements on a CTBT verification regime (including provisions for in-country monitoring and on-site inspection), though briefly considered again in the late 1970s, was effectively postponed for a generation, beginning again on a multilateral basis at the Conference of Disarmament in January 1994 in Geneva, where the work began so many years before.

### **3. Further Comment on Key Technical Issues in Seismic Monitoring**

The work of monitoring underground nuclear explosions using seismological methods can usefully be broken down into the separate steps of signal detection, location of the event, identification, and estimation of yield. All these steps may be studied both for underground tests conducted non-evasively (the practice under the LTBT); and for tests conducted evasively using various methods (some, as yet hypothetical) to reduce or otherwise manipulate signals, with the goal of avoiding detection and/or identification.

Improvements in signal detection came steadily throughout the early years of Project VELA, along with improved methods of event location. The effectiveness and engineering feasibility of the cavity decoupling method of treaty evasion appears to be much more limited than envisioned in 1963. But the early development of methods for event identification began with a setback, and a key method that was indeed successful was discovered too late to have any impact in the period 1958-1963.

The setback, was the early realization that the method of *P*-wave “first motions” was very unreliable, because it is often impossible to be sure if the trace of a seismogram moves up or down at the beginning of the arrival of *P*-waves from either an earthquake or an explosion. (In practice, seismogram signals include a background of noise, and it is common for *P*-waves to emerge as growing oscillations from this background with no clear indication of a first upward or first downward movement.)

The key method of event identification that was successful was based upon use of more than just the *P*-wave. To explain this method, we must first note that several different seismic waves are produced by earthquakes and underground explosions. These waves fall broadly into three types: those which travel through the body of the Earth (i.e. through its deep interior); those which spread out over the surface of the Earth, analogous to the way that ripples disperse over the surface of a pond; and those which are guided along by the outer layer of the Earth (the crust). These three types of waves are referred to as seismic body waves, seismic surface waves, and regional waves. A subdivision of seismic body waves into so-called *P*-waves and *S*-waves has been known since the 19<sup>th</sup> century. As noted above, *P*-waves travel faster than all other

seismic waves, and, though traveling in solid rock, are analogous to sound waves in air or water. *S*-waves (“secundus”) also travel through the Earth’s deep interior, but slower than *P*-waves. They consist of a shearing motion in which particles move at right-angles to the direction the *S*-wave itself is traveling. The typical frequency of a *P*-wave or an *S*-wave, as recorded at teleseismic distances (see Figure 1, the Third Zone), is in the range 0.5 – 5 Hz. Surface waves are also recorded teleseismically, but with much lower frequency, typically around 0.05 Hz. The strongest regional wave, known as the *Lg*-wave, may have frequencies in the range 0.3 – 3 Hz.

A number of seismologists noticed in the early 1960s that the different types of seismic waves were excited to different levels by underground explosions, than was the case with shallow earthquakes. Some of this work had in fact been known for many years, but in the context of studying the signals from quarry blasts, which are typically too small to detect except at regional distances. Thus, a Harvard seismologist, Don Leet, who had specialized in the study of quarry blasting, noted from teleseismic records of underground nuclear explosions that they often lacked any *S*-wave signals even when the *P*-wave was strong [35]. For earthquakes, the *S*-wave is usually much stronger than *P*. The use of what Leet called the “lonesome *P*” discriminant, however, was unreliable, for many nuclear explosion records did in fact include *S*-wave signals. More important, was the discovery that underground nuclear explosions are inefficient, relative to earthquakes, in exciting surface waves. Much of the early work in this field was done at what was then called the Lamont Observatory of Columbia University in New York. For example, Brune and others in 1963 found from the study of more than a 100 earthquakes and 35 explosions that “Most of the earthquakes studied generated surface waves 5 to 10 times greater than the maximum observed for explosions” when the explosions and earthquakes had short-period regional waves of the same size [36]. Lieberman and Pomeroy [37, 38] used traditional methods of measuring the magnitude of teleseismic body waves ( $m_b$ ) and surface waves ( $M_s$ ), and showed for an underground nuclear explosion in the Aleutians that the  $M_s$  value was only 3.9, whereas for an earthquake with the same  $m_b$  as the explosion, the  $M_s$  value would be expected to be about 6.1. They then applied this discrimination method to two seismic events in Southern Algeria and successfully identified them as underground nuclear explosions, because “the surface waves from these events are much smaller than would be expected from most earthquakes of comparable body-wave magnitudes.”

The so-called  $M_s:m_b$  discriminant was clearly successful for shallow seismic events, if they were large enough to give teleseismic body-wave and surface wave signals whose magnitude could be reliably measured. (For deep events, other discriminants could be used.) The method resulted in many efforts over a period of years to improve the ways that  $m_b$  and  $M_s$  are measured, and many efforts to see if the discriminant could be applied reliably at lower magnitudes. Figure 3 shows key results obtained in

1971 (though not released until several years later) for underground nuclear explosions at the Nevada Test Site and earthquakes in Nevada, namely: that the method appeared to be reliable down to below  $m_b$  4; and that the two populations (of explosions and earthquakes) did not appear to merge at low magnitude, so the method could potentially be made to work at even lower magnitudes if signals could be obtained (in particular, surface wave signals from small explosions) [39].

With the growth of the WWSSN in the early 1960s, signal quality was adequate to apply the  $M_s:m_b$  discriminant routinely on a global basis down to  $m_b$  4.5. For example, Sykes and others showed that for events with  $m_b \geq 4.5$ , 90% could be identified as earthquakes based upon their depth being greater than 30 km and/or their location being more than 25 km at sea; and all the remaining 10% could be identified using the  $M_s:m_b$  method [40]. This capability would appear to have been adequate to monitor compliance with a trilateral underground test ban of the type considered in the early 1960s, with a ban on events of magnitude 4.75 and above, although the WWSSN stations would have needed augmentation to improve detections in Eurasia.

*Figure 3.* A robust discriminant, the plot of  $M_s$  against  $m_b$  for earthquakes and underground nuclear explosions. Since magnitude scales are logarithmic, the separation of the two lines by 0.8 magnitude units implies that surface waves from earthquakes are, on average, more than 6 times larger than surface waves from explosions having the same body wave magnitude. From [39].

#### 4. Conclusions

The capability to monitor a CTBT by seismological methods was developed on an accelerated basis in the early 1960s, but was then deemed inadequate, leading apparently

to the need for significant numbers of on-site inspections of suspicious events. In retrospect, we find that monitoring methods turned out to be significantly better than they were typically characterized at the time by key advisors. The greatest improvements in capability were developed in the practical context of monitoring underground nuclear weapons tests following the conclusion of the LTBT – not in the context of earlier arms control negotiations.

The Geneva system of 170 control posts (see Figure 2) was never built, but, on the basis of comparison with other networks, it appears it would have enabled monitoring to be accomplished on a global basis down to  $m_b$  3 rather than  $m_b$  4, about a tenfold improvement over what was stated at the time to be the desired monitoring capability.

## 5. Acknowledgements

Support for over two decades from the Advanced Research Projects Agency, the Air Force Phillips Laboratory, and the Air Force Office of Scientific Research (AFOSR) is acknowledged; with current support being provided by AFOSR through grant F49620-94-1-0057. Note: the views and conclusions here are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government. This is Lamont-Doherty Earth Observatory contribution no. 5301.

## 6. References

1. Kerr, D.M. (1988) in J. Goldblat and D. Cox (eds.), *Nuclear Weapons Tests: Prohibition or Limitation?*, Oxford University Press, New York, p. 43.
2. Magraw, K. (1988) Teller and the Clean Bomb Episode, *Bulletin of the Atomic Scientists*, May issue, p. 32.
3. Gilpin, R. (1962) *American Scientists and Nuclear Weapons Policy*, Princeton University Press, Princeton, NJ.
4. Jacobson, H.K., and E. Stein (1966) *Diplomats, Scientists, and Politicians: The United States and the Nuclear Test Ban Negotiations*, Univ. of Michigan Press.
5. Johnson, G.W. (1985) Underground Nuclear Weapons Testing and Seismology - a Cooperative Effort, in *The VELA Program: A Twenty-Five Year Review of Basic Research*, Defense Advanced Research Projects Agency, p. 10.
6. The New York Times (1958), April 3, p. 1.
7. Conference of Experts to Study the Methods of Detecting Violations of a Possible Agreement on the Suspension of Nuclear Tests (1958) Verbatim Records.
8. Teller, E. (1958), Alternatives for Security, *Foreign Affairs* **36**, p. 204.
9. Murray, T. (1959) East and West Face the Atom, *The New Leader*, June 15, pp. 10-14.
10. US Congress (1959) Senate Committee on Foreign Relations, Subcommittee on Disarmament, Hearings: Geneva Test Ban Negotiations, 86th Congress, 1st Session, p. 29.
11. Ringdal, F. (1985) Study of magnitudes, seismicity and earthquake detectability using a global network, in *The VELA Program: A Twenty-Five Year Review of Basic Research*, Defense Advanced Research Projects Agency, p. 611.

12. Romney, C. (1960) Detection of Underground Explosions, in *Project Vela, Proceedings of a Symposium*, October, pp. 39-75.
13. Murphy, J. (1981) P-wave coupling of underground explosions in various geologic media, in E.S. Husebye and S. Mykkeltveit (eds.), *Identification of Seismic Sources – Earthquake or Explosion*, D. Reidel, Dordrecht, pp. 201-205.
14. Ringdal, F., P.D. Marshall and R. Alewine (1992) Seismic yield determination of Soviet underground nuclear explosions at the Shagan River test site, *Geophysical Journal International* **109**, 65-77.
15. US Department of Defense (1959) Press Release, Jan 16.
16. Latter, A., R. LeLevier, E. Martinelli, W. McMillan (1959) A Method of Concealing Underground Nuclear Explosions, RAND Corporation, Mar. 30. Subsequently published (1961) *Journal of Geophysical Research* **66**, 943-946.
17. Street, K. (1959) Need for High Explosive and Nuclear Tests for Research Program, *Report of the Berkner Panel*, p. 54.
18. Report of the Berkner Panel (1959), p. 15.
19. Project Vela, Proceedings of a Symposium (1960), October.
20. Bethe, H. (1960) The Case for Ending Atomic Testing, *The Atlantic Monthly* **206**, p. 48.
21. Conference on the Discontinuance of Nuclear Weapon Tests, Geneva (1958-62) Verbatim Records, section 188, p. 13.
22. Latter, A. (1960) Decoupling of underground explosions, in *Project Vela, Proceedings of a Symposium*, October, p. 180.
23. Kistiakowsky, G.B. (1976) *A Scientist at the White House*, Harvard University Press, Cambridge.
24. Romney, C. (1962) US Congress, Joint Committee on Atomic Energy, Hearings: Developments in the Field of Detection and Identification of Nuclear Explosions (Project Vela) and their Relationship to Test Ban Negotiations, 87th Congress, 1st Session, pp. 123-4.
25. Seaborg, G.T. (1981) *Kennedy Khrushchev and the Test Ban*, Univ. of California Press, Berkeley.
26. Khalturin, V., T. Rautian, and P.G. Richards (1994), A study of small explosions and earthquakes during 1961-1989 near the Semipalatinsk test site, Kazakhstan, Technical Report to AFOSR, 1994.
27. Båth, M. (1962) Seismic records of explosions - especially nuclear explosions: Part III, Forsvarets Forskningsanstalt (Swedish Defense Research Establishment), FOA report 4, A 4270-4721, December, pp. 60-63.
28. Agnew, H. (1987) personal communication to PGR.
29. Bethe, H. (1962) Disarmament and Strategy, *Bulletin of the Atomic Scientists* **18**, 14-22.
30. US Congress (1963) Joint Committee on Atomic Energy, Hearings: Developments in Technical Capabilities for Detecting and Identifying Nuclear Weapons Tests, 88th Congress, 1st Session.
31. Foster, W. (1962) U.S. Congress, Disarmament Subcommittee Hearings: Renewed Geneva Negotiations, 87th Congress, 2nd Session, 25 July.
32. US Congress (1963) Senate Committee on Foreign Relations, Hearings: Test Ban Negotiations and Disarmament, 88th Congress, 1st Session, 1963.
33. US Congress (1962) Senate Armed Services Committee, Preparedness Investigation Subcommittee, Hearings: Arms Control and Disarmament, 87th Congress, 2nd Session, p. 13.
34. Teller, E. (1963) US Congress, Senate Committee on Foreign Relations, Hearings: Nuclear Test Ban Treaty, 88th Congress, 1st Session, p. 428.
35. Leet, D. (1962) The detection of underground explosions, *Scientific American* **206**, 55-59.
36. Brune, J., A. Espinosa, and J. Oliver (1963) Relative excitation of surface waves by earthquakes and underground explosions in the California-Nevada region, *Journal of Geophysical Research* **68**, June 1, 3501-3513.
37. Liebermann, R.C., C.-Y. King, J.N. Brune, and P.W. Pomeroy (1966) Excitation of surface waves by the underground nuclear explosion Longshot, *Journal of Geophysical Research* **71**, 4333-4339.
38. Liebermann, R.C., and P.W. Pomeroy (1967) Excitation of surface waves by events in Southern Algeria, *Science* **156**, 1098-1100.
39. Lambert, D.G., and S.S. Alexander (1971) Relationship of body and surface wave magnitudes for small earthquakes and explosions, SDL Report 245, Teledyne Geotech, Alexandria, Virginia.
40. Sykes, L.R., J.F. Evernden, and I. Cifuentes (1983) Seismic methods for verifying nuclear test bans, in D.W. Hafemeister and D. Schroerer, *Physics, Technology and the Nuclear Arms Race*,



Liebermann and Pomeroy/quotes, reference

Layout of Figures 1, 2 & 3

Permission from Zavales

AFOSR grant no.

LDEO Contribution no.

Ask Anton, do they go with my camera-ready, or do they edit e.g. from a disc?

Get lat/long of 170 control posts/send to John Claassen/request detection capability

Make Figures for talk, e.g. mag-yield curves, no. of events/year, cf. modern estimates.