Uncertainty in ASTM Test Methods and ISO/IEC 17025

Verification and Validity

By Neil Ullman

Q How will the changes in ISO/IEC 17025 impact how uncertainty is treated in ASTM?

A A new version of ISO/IEC 17025, general requirements for the competence of testing and calibration laboratories, was adopted in late 2017 (a previous version was issued in 2005). The organization of the standard has been revised, and some new topics and approaches have been incorporated. Although the primary focus seems to follow the main principles of the earlier versions, there are changes that may be relevant to ASTM standards writers and users.

A major new portion of the standard addresses verification: "provision of objective evidence that a given item fulfills specified requirements." Part of this is "confirmation that a 'target measurement uncertainty' can be met."

This addition opens up the possibility that an ASTM International standard specification might provide a target uncertainty. Thus, there may be an important sharing of responsibility between specification writers and test method developers for appropriate approaches to meet a desired uncertainty. Of course, this also opens up a major concern about understanding the nature of variation that is both inherent in the measurement process as well as the materials being evaluated, especially regarding the roles of the consumer and the standards community.

In ISO/IEC 17025:2005 there was no discussion of verification. There was an extensive section on the validation of methods, especially for those that are non-standard, laboratory-developed, or outside the normal scope. Now, with the definitions of verification and validation in the VIM (*International Vocabulary of Metrology*, 3rd edition), the emphasis has shifted to verification, with validation being considered as a special case of verification.

We also see the term "validity" and a whole section on "ensuring the validity of results." The term itself is not defined. The primary way validity is required is that there shall be a procedure for monitoring the quality of the test and measurement results.

The key term here is monitoring. This very clearly presents the need to pay attention within the laboratory to the quality control of test procedures. An important ASTM standard, the practice for estimating and monitoring the uncertainty of test results of a test method using control chart techniques (E2554), provides a useful approach both for obtaining an estimate of uncertainty through actual test results, as well as providing an ongoing ability to monitor the uncertainty (change in precision or bias) over time. This uses the within-laboratory long-term intermediate precision as an estimate of uncertainty for the laboratory. Test method performance can also be monitored using the practice for the use of control charts in statistical process control (E2587).

ISO/IEC 17025 evolved from a primary function of evaluating metrological measurements – to provide a procedure to calibrate the instruments in a traceable system back to a national reference standard. The focus was on the small set of primary measures: length, time, mass, electric current, temperature, amount of substance, and luminous intensity. The idea of measurement uncertainty emerged in this context, and attempts have been made to expand this notion to equipment and test methods that have little such traceability.

The latest version of the standard has continued this effort and has suggested that under certain circumstances one could use information from an interlaboratory study presented in a published standard such as an ASTM International test method to satisfy requirements for presenting estimates of uncertainty.

Among the concerns with this approach is first the fact that what is reported as repeatability, the within-lab short-term estimate of variability, is averaged over all laboratories that participated in the study. Some would have had worse levels of variation than others and no lab can assume it automatically would perform like any other lab. The second item reported, reproducibility, is partially the result of relative biases among the participating labs. For these reasons, Section A22 of *Form and Style for ASTM Standards* interprets uncertainty to be exclusively the responsibility of a given lab and that estimates from an interlaboratory study do not necessarily describe how a given method would behave in a particular lab.

One way of viewing the difference in associating uncertainty in metrological measurements versus testing is how calibration is considered. For metrological measurements, the evaluation of uncertainty applies to the instrument through the calibration process that does not necessarily address the myriad of factors that can influence the measurement. It is, however, expected that the lab shall incorporate those possible influences through the development of an uncertainty budget. To some extent such contributions to uncertainty may be estimated through actual measurements. For some discussion, see for example, the guide for reporting uncertainty of test results and use of the term measurement uncertainty in ASTM test methods (E2655).

Calibration involving a traceable system does not apply to many ASTM test methods, since there are very rarely any true reference standards. In most situations the equipment is not "standardized," and there is very minimal theoretical understanding of the actual test method process.

Perhaps even more important is that the individual operators and environments are not formally a part of the process. As much as we would like to predict and model the way a test method is undertaken, the reality is that it is essential to recognize that each test environment must develop and monitor its own practices and use the facts so accumulated. The guide for statistical procedures to use in developing and applying test methods (E1488) emphasizes that "It is neither appropriate for, nor the responsibility of the test method to provide values of uncertainty that a user should use as their estimate of uncertainty." This argument is further reiterated in Section A22 of *Form and Style*.

A recent series of studies involved a new standard for use of various Vernier calipers to measure the width of sheet materials. In one case a series of measurements were taken on a fixed set of different size materials, using a single set of calipers, by a number of operators in the same facility. Significant differences in the results by operator were revealed on all materials and instruments. Other experiments with the same materials involved multiple sites where similar sets of calipers were used, and again variations among labs and operators were unpredictable. In addition, some labs had consistently precise results and others much greater variation.

This was with instruments that are to be calibrated with valid reference materials such as gage blocks. What happens with those ASTM methods that deal with the many types of materials, such as asphalt, fuel oil, paint, and rubber, for which it is difficult to find uniform, non-changing standard materials that are then tested with varying types or models of equipment?

ASTM recommends that when an interlaboratory study is conducted it should consider following the practice for conducting an ILS to determine the precision of a test method (E691). This standard stipulates that participation in an ILS should only be taken by "qualified" laboratories (see Section 9.3), which implies that the operators are trained and familiar with the test method. The requirement for uncertainty estimation in ISO/IEC 17025 is with us to stay. Perhaps one further way of qualifying laboratories would be based on their ability to estimate their uncertainty.

One option would be to have ILS participating laboratories first establish estimates of uncertainty before executing any of their ILS testing. The estimates of uncertainty could then also be incorporated in the study, and the final reporting might provide information on how well labs actually perform their uncertainty analyses. This also would require the test method developers to first provide adequate recommendations for establishing an estimate of uncertainty and then examine how well those recommendations or other methods performed in actual testing. This would be equally useful for any proposals for new or revised test methods to have such results incorporated in the preliminary submission of standards. A new subcommittee on metrology (E11.50) has recently been established and will consider developing additional guidance documents.

Neil Ullman, a consultant with Ullman Associates, Florham Park, New Jersey, is a retired professor of mathematics and mechanical technology from the County College of Morris, and a past chairman of the committee on quality and statistics (E11), where he is currently a member at large on the executive subcommittee and co-chair of the subcommittee on metrology (E11.50). He is a fellow of ASTM International and the American Society for Quality.

John Carson, Ph.D., of P&J Carson Consulting LLC, Findlay, Ohio, is the Data Points column coordinator. He is chairman of the subcommittee on statistical quality control (E11.30), part of the committee on quality and statistics (E11), and a member of the committee on environmental assessment, risk management, and corrective action (E50).