Point of View

Highway Quality Assurance

RICHARD M. WEED

Are We Living Up to the NQI?

The network of roads and bridges linking the suppliers of goods and services with their customers is one of the nation's most valuable assets. National economic well-being is dependent on the condition of the highway system, which in turn is inexorably linked to the quality of design and construction. To control the quality of construction, highway agencies have developed elaborate quality assurance programs based on statistical sampling and acceptance procedures and designed to ensure that work is done in accordance with plans and specifications. Yet while statistical quality assurance is currently employed by more than three-quarters of the states, many procedures in use today were developed long before statistical science was a requirement in the typical engineering curriculum. Consequently, many existing specifications for statistical quality assurance have never been thoroughly analyzed to confirm that they will perform as intended.

In response to a growing concern that quality assurance practices were not providing the desired degree of highway quality, the Federal Highway Administration, the American Association of State Highway and Transportation Officials, and several industry associations joined forces in 1992 to form a unique partnership called the National Quality Initiative (NQI). Its stated goal was to promote the improvement of highway product quality as one of the keys to national economic competitiveness. The long-range plan of the NQI defines a broad series of objectives and activities designed to raise the level of consciousness on quality issues. These activities are focused heavily on education, training, and information exchange in the belief that knowledge and understanding are ultimately the best motivators. This effort is consistent with and will be aided by the current strong interest in total quality management, which emphasizes continuous improvement, data-driven decision making, and the notion that quality is everyone's responsibility.

Not all quality assurance programs are effective, however, nor are all acceptance procedures fair to both the contractor and the highway agency. Statistical quality assurance is one of the most useful tools a highway agency has at its disposal, but only if it is used correctly. Are we really performing adequately as a profession when we use acceptance procedures that require a 20 percent greater sampling and testing effort to accomplish the same result? Or data comparison procedures that cannot reliably detect a difference in concrete compressive strength as large as 1000 psi? Or pay schedules that pay well below 100 percent, on average, even though the contractor has provided exactly the level of quality specified by the contract documents? Or acceptance procedures that give greater reward to the contractor that has provided lesser quality? Or specifications that are so weak they may allow seriously defective work to slip by undetected? Or voluntary consensus standards that rely exclusively on the balloting process to ensure their validity? Unfortunately, these are not hypothetical examples; they are well documented and are typical of highway quality assurance practices in use today.[1]

Even national guide specifications, which one might think should be held to the highest standards of technical rigor, are prone to the above shortcomings. Consider the following procedure, the purpose of which is to determine whether a set of test results obtained by a contractor is consistent with a single test performed by a state agency. Based on the range (R) of the contractor's data set, the test requirement is given by equation (1), and the appropriate range coefficient (C) is obtained from Table 1.

[1] Two recent reports that detail these concerns are *Managing Quality: Time for a National Policy,* NJDOT Research Report 96-004-7490, June 1996, and *Highway Quality Assurance: The Need for Science- and Knowledge-Based Decision Making,* NJDOT Research Report 97-001-7490, March 1997. Both reports may be obtained from NJDOT or the National Technical Information Service.

The author is supervising statistical engineer, New Jersey Department of Transportation.

$$\overline{X} - C * R \leq X \leq \overline{X} + C * R \quad (1)$$

where

$\overline{X}$ = average of several quality control tests performed by the contractor,
R = range of the quality control tests,
C = coefficient by which the range is multiplied, and
X = single quality assurance test performed by the transportation agency.

To evaluate such a decision procedure, three key questions must be answered:

- What is the probability of falsely detecting a difference between the two data sets when there truly is no difference?
- What is the probability of correctly detecting a difference between the two data sets when such a difference truly exists? For example, when considering the expected performance of concrete pavement or structures, a deficiency in compressive strength of 6890 kilopascals (kPa) (1000 psi) might be associated with a significant loss in service life or performance. Therefore, it would be very beneficial to know the likelihood of correctly detecting a difference of this magnitude.
- Similarly, it would be useful to know how large the true difference between two populations must be if there is to be a reasonably high probability of detecting that difference. For example, again using concrete compressive strength, what must the true difference in population means be if there is to be a 0.90 probability that the difference will be detected?

To answer these questions, an operating characteristics curve analysis must be performed. Table 2 shows results for the case in which there are N = 5 contractor tests. The true difference between the means of the two populations is given in standard deviation units so the results can be generalized to a wide variety of applications.

TABLE 1 Range Coefficients for Use with Single-Test Decision Procedure

Number of Contractor's Tests (N)	*Range Coefficient (C)*
5	1.61
6	1.33
7	1.17
8	1.05
9	0.97
10	0.91

TABLE 2 Operating Characteristics of Single-Test Decision Procedure

Difference in Means (SD Units)	*Probability of Detecting Difference*
0.0	0.02
1.0	0.05
2.0	0.17
3.0	0.35
4.0	0.57
5.0	0.76
6.0	0.90

The answer to the first question above can be read from the first row in Table 2. A risk of 0.02 of falsely detecting a difference when none exists would probably be tolerable for most applications. The answers to the second and third questions reveal a serious shortcoming, however. For example, depending on the length of time over which the data are gathered, the standard deviation associated with concrete compressive strength could be as large as 3445 kPa (500 psi). To answer the second question, a difference of 6890 kPa (1000 psi) would be equivalent to 2.0 standard deviation units, and, based on the results in Table 2, the probability of detecting a difference this large when it truly exists is only about 0.17. To answer the third question, Table 2 indicates that a true difference of 6.0 standard deviation units is necessary for there to be a 0.90 probability of detection. Therefore, the actual difference in concrete compressive strength must be 6.0 x 3445 = 20 670 kPa (3000 psi) if there is to be a reasonably high probability of detection.

Figure 1 illustrates a typical application of this procedure in which the sample mean ($\overline{X}$) falls close to the population mean, and the five individual samples have about the same dispersion as the parent population. The figure reveals how this procedure readily leads to the conclusion that there is no difference between the two data sets, even when the two populations are completely separated. Obviously, there are not many engineering applications for which such a weak procedure could be considered practically useful. In fact, the procedure could provide a false sense of security by continually showing no significant difference between the data sets when the true differences are actually quite large. In many cases, seriously defective work could be accepted without the agency being aware of it.

The above example demonstrates the risk inherent in attempting to make decisions based on an insufficient amount of data and is but one of many examples that could be offered. The lesson to be learned is simply that, in order to realize the many benefits statistical science has to offer and avoid potentially costly mistakes, these procedures must be applied with knowledge and understanding.

But are these matters of real concern? After all, the highway system appears to be performing reasonably well, and deficient quality probably is not going to result in any fatalities. A little reflection, however, will suggest that such a complacent attitude may be extremely shortsighted. As noted earlier, the nation's economic well-being is strongly linked to the condition of our transportation system, and nationwide condition surveys have the disquieting habit of declaring substantial portions of the system to be in critical states of disrepair. The replacement value of our network of roads and bridges has been estimated to be in the trillions of dollars. No stretch of the imagination is required to see that laxness on quality issues today could put great strains on future transportation budgets.

The costs of poor practices go beyond simple economics. Poorly developed programs do little but breed disrespect, distrust, a generally adversarial atmosphere, and a determination on the part of the construction industry to oppose the programs. In a discipline dedicated to excellence, it is inappropriate to tolerate either specifications or consensus standards that are far from excellent. If demands for excellence are to be placed on the construction industry, transportation agencies must place the same demands on themselves in developing the standards and specifications that govern the work. Critical decision-making processes in highway construction—processes that affect the acceptance or rejection of work performed by the construction industry or the level of compensation for such work—must be based on sound engineering and mathematical principles. Only in this way will it be possible to develop specifications and acceptance procedures that allow quality-conscious contractors to operate at a fair profit.

In recognition of the seriousness of this problem, two committees of the Transportation Research Board—Management of Quality Assurance, and Statistical Methodology and Statistical Computer Software in Transportation Research—jointly submitted a research needs statement for the annually conducted national ballot to identify critically needed pooled-fund studies. This problem statement, entitled "Optimal Acceptance Procedures for Statistical Construction Specifications," is somewhat broader than its title implies and is aimed at putting statistical quality assurance on a sound scientific and mathematical footing. The goal is to demystify basic quality assurance technology and make it rationally comprehensible to its many users.

This problem statement was one of 46 that made the final ballot and that were rated by 39 state highway agencies. It was the clear front runner in the balloting. As a result, a formal study has been approved, funding commitments have been obtained from the many states that supported the concept, and a consultant to do the work is currently being sought.

This study will almost certainly lead to significant improvements in quality assurance practices over the long term, but there are other steps conscientious professionals can take in the meantime. Perhaps the most critical need is for transportation leaders to recognize that the scientific and technical aspects of quality assurance are not minor details that can be taken for granted. A conscious effort must be made to seek out the best methods and use them correctly. If this simple change in mind-set could be accomplished, many of the other problems associated with quality assurance would soon disappear.

Highway administrators could then be confident that properly designed quality assurance procedures would benefit both the organization and the public it serves. The construction industry would be less resistant, and possibly even supportive, because technically sound practices can offer conscientious contractors a competitive advantage. Voluntary consensus groups might begin to seek out the necessary expertise as part of the standards-writing process, and require independent reviews to ensure that any guidelines offered to the rest of

continued on page 36

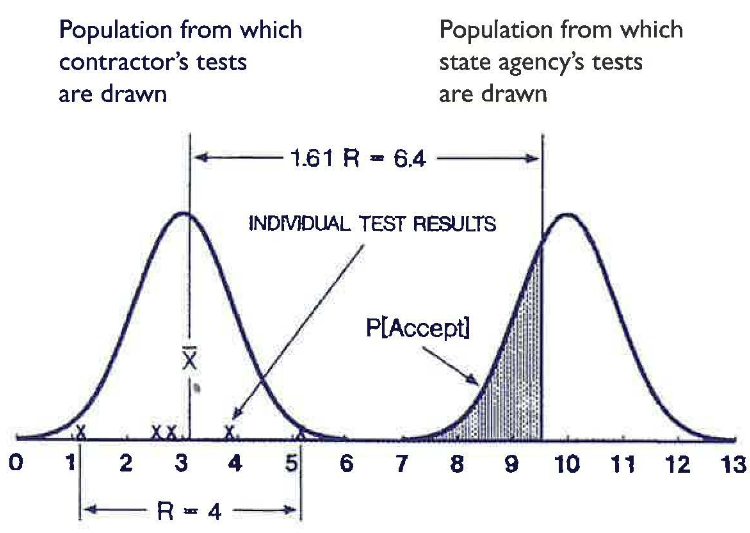

FIGURE 1 Illustration of weakness of decision procedures based on a single test.

Profiles

Evans *continued from page 24*

up." After completing a bachelor's degree in physics in 1960, he continued his education at the University of Oxford in England, receiving a doctor of philosophy degree in 1965. His interest in applying physics and mathematics to experimental and theoretical investigations of traffic phenomena was inspired by the late Dr. Robert Herman, formerly Head of the Theoretical Physics Department of General Motors Research Laboratories, whom he met in 1966 during post-doctoral work at the Pure Physics Division of the National Research Council of Canada.

Throughout his career, Evans has retained his sense of humor. For example, he remembers that a general physics text used by his elder son while a freshmen at the Massachusetts Institute of Technology contained a reference to one of his traffic safety publications. Commenting on this he remarked, "If I had remained in physics, such a text would hardly have included my name; to be included in an elementary physics text, it generally helps to be dead for at least 100 years."

Gulden *continued from page 25*

has authored or coauthored a number of research papers and reports. In 1981 he received TRB's D. Grant Mickle Award for a paper entitled "Pavement Restoration Measures To Precede Joint Resealing." Currently the TRB representative for the state of Georgia, he is a member of the TRB Expert Task Group on Long-Term Pavement Performance Automated Distress Identification, the ETG for Pavement Preservation, and the National Cooperative Highway Research Program Panel "10–50 Strategies for Rehabilitating Rigid Pavements Subjected to High-Traffic Volumes." He is a team leader on the SHRP Implementation Lead State Team for Pavement Preservation and is a member of the Federal Highway Administration ETG SP205, Quality Concrete Pavement Rehabilitation and Preservation. In addition, he serves on the American Association of State Highway and Transportation Officials Region II Research Advisory Committee and on the AASHTO Subcommittee on Materials. He is also a member of the Highway Innovative Technology Evaluation Center panel for PH-100 Pothole Repair Technology.

Highway Quality Assurance

continued from page 9

the profession would be mathematically and scientifically sound. And by advocating the correct use of the most effective methods, it may ultimately be possible to establish the consistent use of sound practices nationwide. This would clearly be a win-win situation for all concerned.

To bring all this about, it is necessary to create an environment that welcomes and encourages self-criticism and continuous improvement, one in which excellence is regarded as the norm, and pride of workmanship is once again a primary motivating force. The necessary technology exists, as does the talent to use that technology, but the processes that would ensure their proper use have yet to be put in place. As a profession we are capable of doing much better, and in an increasingly competitive technological age, it would appear unwise to aim for anything less than our best.