TRANSPORTATION RESEARCH

RECORD

No. 1340

Materials and Construction

Managing Quality: A National Perspective

A peer-reviewed publication of the Transportation Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
WASHINGTON, D.C. 1992

Transportation Research Record 1340
Price: $18.00

Subscriber Category
IIIB materials and construction

TRB Publications Staff
Director of Publications: Nancy A. Ackerman
Senior Editor: Naomi C. Kassabian
Associate Editor: Alison G. Tobias
Assistant Editors: Luanne Crayton, Norman Solomon
Production Coordinator: Karen W. McClain
Office Manager: Phyllis D. Barber
Production Assistant: Betty L. Hawkins

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data
National Research Council. Transportation Research Board.

Managing quality : a national perspective.
p. cm.—(Transportation research record, ISSN 0361-1981 ; no. 1340)
"A peer-reviewed publication of the Transportation Research Board."
ISBN 0-309-05204-1
1. Highway engineering—Quality control. 2. Roads—Design and construction—Quality control. 3. Pavements—Quality control. I. National Research Council (U.S.). Transportation Research Board. II. Series: Transportation research record ; 1340.
TE7.H5 no. 1340
[TE153]
388 s—dc20
[625.7'0685] 92-22591
CIP

Sponsorship of Transportation Research Record 1340

GROUP 2—DESIGN AND CONSTRUCTION OF TRANSPORTATION FACILITIES
Chairman: Charles T. Edson, New Jersey Department of Transportation

Construction Section
Chairman: Donn E. Hancher, Texas A&M University

Committee on Management of Quality Assurance
Chairman: Orrin Riley, Orrin Riley PE PC
Secretary: Richard M. Weed, New Jersey Department of Transportation
Paul E. Benson, Timothy D. Bowen, E. J. Breckwoldt, Gerard Brunschwig, James L. Burati, Jr., Humberto Castedo, C. David Church, Ronald J. Cominsky, Michael I. Darter, Ervin L. Dukatz, Jr., Robert P. Elliott, Ralph D. Ellis, Jr., Jim Gee, Myron Geller, Charles S. Hughes, Rosemary M. Ingram, Charles A. Kline, Peter A. Kopac, M. Lee Powell III, Randell C. Riley, Carlos Rosenberger, James M. Shilstone, Sr., Peter A. Spellerberg

Frederick D. Hejl, Transportation Research Board staff

The organizational units, officers, and members are as of December 31, 1991.

Transportation Research Record 1340

Contents

Foreword

The concept of total quality management (TQM) has challenged all areas of endeavor in recent years, certainly including the transportation sector. The concept requires a multidirectional examination of all functions and processes whether in planning, administration, design, construction, or maintenance in order to create a reliable product.

Meanwhile the TRB Committee on Management of Quality Assurance has been exploring and promulgating the benefits of end-result specifications, especially those rooted in mathematically sound, statistically based acceptance procedures. The committee has also given voice to the promising results of incentive and disincentive clauses in improving quality.

Beginning in 1990, the committee saw the need to integrate its interest in improved specifications and tangible incentives into the larger fabric of TQM. Thus a special session was held at the TRB Annual Meeting in Washington in January 1992 to address that need. The papers presented at that session are published in this Record.

Statisticians Hunter and Pendleton, in the opening paper, explain the different roles of statistical science, from serving as a "number librarian" to screening out the "noise" in data. They conclude that the analytical power of statistics must be embraced if transportation technology is to keep pace with science.

Afferton et al. follow with a four-part paper remarkable in both depth of detail and profundity of persuasion. They issue a "call to action" by pointing out in Part 1 the shortcomings in current procedures and in Part 2 obstacles and the way to overcome them. Part 3 is a virtual handbook for writing reliable specifications. In Part 4 the authors set forth their strategy for winning acceptance of the need to change.

Oswald and Burati cast a practiced eye on the specific needs inherent in public construction and analyze the efforts of certain federal agencies to practice TQM. They conclude that TQM will work, but only if the attitudes of the people involved are accommodating to its success.

Shilstone uses a contrariant approach by advocating analysis in depth of pavements that perform well and not just those that perform poorly. He then extends the end-result hypothesis by urging that it can be applied to the end-result of the pavement (its durability) and not just the end-result of the construction process.

Tuggle closes out the issue by describing how the Federal Highway Administration is developing a demonstration project on quality management. Moreover, that agency is taking the lead in establishing a dialogue with the states and private industry in what is called a "joint policy initiative" to improve quality in its domain.

Viewed separately, these five papers present a thoughtful, sometimes provocative insight into the problems faced and the solutions proposed. Taken as a whole, this compendium can be seen to deliver valuable evidence that the "Quest for Quality" is well under way.

On the Importance of Statistical Science in Transportation

J. STUART HUNTER AND OLGA J. PENDLETON

Sound statistical science must be combined with modern technology if the United States is to meet its transportation quality objectives. Success will require standardized measurements, the collection of information-laden data, resourceful data analysis, and the planning for new data. These quantitative arts must be combined with methods for problem solving and decision making under uncertainty. Statistical science thus joins the engineering sciences in the never-ending pursuit of transportation quality.

Total quality management (TQM) in transportation can be viewed as the concurrent application of engineering, economics, organizational behavior, and public service. Providing TQM in transportation is complicated by the myriad political constituencies, vast geography, and the many facets of transportation—highway, rail, water, and air—in the United States. Nevertheless, certain persistent strands of TQM can be identified whatever the arena of application. Transportation executives lead and incorporate quality values into everyday management; vendors and builders work to provide products that go beyond the mere meeting of specifications. Never-ending improvement coupled with customer satisfaction is the leading objective of TQM for manager and vendor alike (*1,2*). But like the warp and woof of a fine fabric, these interpersonal and human behavior strands must be coupled with the strong cords of technology. Quality transportation *technology* must accompany quality transportation *management*.

The 71st Annual Meeting of the Transportation Research Board reflected the vast range of competencies that together define "transportation technology." One important technical competence is statistical science, its applications often so ubiquitous that many technologists and managers seem almost unaware of its existence. The days wherein transportation leadership can assume this casual attitude toward the art of good statistical practice must pass. Too much is at stake.

To elucidate the importance of modern statistics on behalf of transportation technology, we can begin with the problems of simply recording data. After all, being a good number librarian is a commonplace view of a statistician. Certainly it is important to count and record special occurrences, be they home runs or traffic fatalities. And it is similarly wise to have someone record continuous measurements such as the axle weight of trucks or the compaction of bituminous concrete. Of course, a good number librarian will keep neat files and will be thorough and quick to provide data on demand. But one observes that it is a rare computer software program that cannot meet these humble requirements. The science of statistics on behalf of transportation technology goes far beyond number librarianship.

Statistical science influences data recording by noting that information and data are not identical. A nonstatistical mind confuses the two. What is not realized is that information lies within data, much as ore within a matrix of rock. The statistical art is to create and record *information-laden* data. Attention turns now to what is being monitored or measured, how the data are being collected and recorded, and the purpose of the entire exercise. Furthermore, data are rarely cheap. Statistical problems abound.

Consider the problems of measurements, the "coins," the items of value, regularly interchanged by scientists and engineers (and politicians too). As an example, consider measuring pavement thickness. How much faith can one place in this simple evaluation? Much depends on the completeness of the definition of what is meant by "pavement," a description of how thickness will be measured, with what instruments, where on the pavement, and how many samples across what time and space. Thickness measurements may ultimately be used to estimate pavement stiffness, a subject of much concern in the current Strategic Highway Research Program (SHRP). Unfortunately, in today's world of transportation technology, different groups often hew to different measurement standards. Further, even when a single national transportation measurement standard exists, little or no surveillance or traceability mechanism is available. Frankly, a measure of pavement stiffness made in Boston would not be trusted in Seattle. And if the measurement were not trusted, how would one then make valid comparisons between suggested pavement improvement methods made in the two areas?

It is not reasonable to ask that every measurement made at every time and place have the same integrity as those guaranteed by the National Institute of Standards and Technology (NIST). But it is reasonable to ask those responsible for spending large sums on transportation projects to become more aware of the value of good measurement. Regrettably, most of the measurement systems mandated by the various U.S. transportation agencies are inadequately controlled. The societal costs of these poor measurements are huge (*3*).

The great engineering effort called the Tower of Babel came to a gradual halt when the builders found that they could no longer communicate. The present great engineering accomplishments of modern transportation are there for all to see. Unfortunately, symptoms of faulty technical communication (poor data) dim their present health and future growth. Defining transportation measurement processes, writing product and performance specifications, and setting up proper acceptance procedures require a common technical language.

J. S. Hunter, 503 Lake Drive, Princeton, N.J. 08540-5639. O. J. Pendleton, Texas A&M University, College Station, Tex. 77843-3135.

When one is dealing with data and its associated activities, statistics is the lingua franca.

Alert engineers and managers also recognize that an intrinsic variability, *noise*, accompanies all physical measurements. The variance of measurements (or equivalently its positive square root, the standard deviation) measures data noise and should accompany every reported statistic or set of data. The enhancement of quality requires the ability to detect signals in the presence of noise. For example, statistical quality control is the science of making decisions under uncertainty (accepting or rejecting products) using noisy data. As noted by Afferton et al. in another paper in this Record, successful statistical quality assurance (SQA) programs require carefully written protocols along with estimates of measurement and sampling variance, the statistician's technical term for noise. Awareness of statistical science begins with the knowledge that data variance is a natural phenomenon that is measurable and an essential component of decision making.

But the real leap to a consciousness of the importance of statistical science comes after data are in hand. Data analysis begins. Are patterns evident in the historical record that graphical displays might expose? What descriptive statistics (averages, percentages, ratios, physical measures) should one construct from the data? How might the data be partitioned to display differences or similarities? Using graphical techniques, much statistical analysis can be performed quite simply (*4,5*). Transportation managers should be alert to, and insist upon, informative graphics and learn the pitfalls of graphical misrepresentation (*6–8*).

Analysis then turns to inference. The key question becomes, "What can be inferred in the dark of these data?" What associations exist between the observed variables? What forecasts are reasonable? Are there mathematical or physical models evident, or partially discernible, that might serve to increase one's basic knowledge? How well are the parameters of interest estimated (in statistical terms, what are the standard errors of the estimates)? And finally, what about measures of uncertainty itself?

Concerning the fitting of models, some caution should be exercised when models are constructed from data using only statistical principles. Most statistical models relating cause-and-effect variables are *empirical*, often mere polynomial relationships between *x*'s and *y*'s. Fitting "regression" models is easy in these days of modern computers and statistical software. But the world of transportation science often employs mathematical models that go far beyond these simple statistical constructions—models that entrain the laws of engineering and physics, nonlinear in their parameters and often dynamic. Successful inference requires a blending of engineering knowledge and statistical practice. Enlightened empiricism accompanies good science.

We come finally to the creation of new data, information-laden data to provide answers to questions. A sampling program may be needed to determine the acceptability of purchased materials or the completion of a contract. A survey may be required to measure the public's concern over a proposed project. Special information may be needed in a laboratory or field test to compare different construction methods or items. Statistical science has made profoundly important contributions to the art of planning for data (*9*). Good data speed understanding; poor data become the pollutants to rational planning and decision making.

On reflection, one discovers that the science called "statistics" is the formalization of the scientific method; it is the use of data within the learning process applied to problem solving. Statistical science begins with the arts of amassing data, provides methods for data analysis leading in turn to inferences elucidating fundamental principles, and finally forecasts future performance, requiring confirmation and the need for new data. This learning process repeats the cycle of data gathering, analysis, inference, and forecasting until a state of knowledge is acquired sufficient to solve the problem at hand. Deming's insistence on "never ending improvement" reflects the essential need to increase knowledge on behalf of quality, to *learn*.

Managers and engineers in transportation and in all the applied sciences employ data daily to solve problems and thus daily "do statistics." What is generally lacking is an awareness that methods for problem solving and decision making under uncertainty are themselves a science. Its name is "statistics." Someone once said that problems are most quickly solved using the highest level of language possible (*10*). The challenge then is to make the scientific method, the learning process, the logic and language of statistics more apparent and less subliminal. This will only be accomplished by education. A good place to begin would be with the simple art of statistics.

REFERENCES

1. W. E. Deming. *Out of the Crisis*. MIT Press, Cambridge, Mass., 1986.
2. J. M. Juran. *Juran on Leadership for Quality*. The Free Press, Macmillan, New York, 1989.
3. J. S. Hunter. The National System of Scientific Measurement. *Science*, Vol. 210, 1980, pp. 869–874.
4. J. W. Tukey. *Exploratory Data Analysis*. Addison-Wesley, Reading, Mass., 1977.
5. J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Tukey. *Graphical Methods for Data Analysis*. Wadsworth, Belmont, Calif., 1983.
6. E. R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, Conn., 1983.
7. E. R. Tufte. *Envisioning Information*. Graphics Press, Cheshire, Conn., 1990.
8. H. Wainer. How to Display Data Badly. *American Statistician*, Vol. 38, No. 2, 1984, pp. 137–147.
9. G. E. P. Box, W. G. Hunter, and J. S. Hunter. *Statistics for Experimenters*. Wiley, New York, 1978.
10. C. H. Kepner and B. B. Tregoe. *The New Rational Manager*. Princeton Research Press, Princeton, N.J., 1981.

Managing Quality: Time for a National Policy

KENNETH C. AFFERTON, JACK FREIDENRICH, AND RICHARD M. WEED

One of the nation's most valuable assets is the highway system; U.S. economic well-being depends strongly on the condition of the country's roads and bridges. Any means by which the system can be more effectively constructed and maintained warrants thoughtful consideration. Statistical quality assurance—currently in use or under development in approximately three-fourths of the states—has proven to be a very effective tool to encourage high-quality construction. However, although statistical specification writing must now be recognized as a thoroughly scientific activity, there is great disparity in the applications from state to state and many current practices and published standards are far from optimal. Part 1 of this paper stresses the need for sweeping reforms and suggests that the time is overdue for the establishment of a uniform and thorough national policy on transportation quality assurance. Part 2 describes a variety of obstacles—technical, managerial, political, and cultural—that must be overcome if such a transformation is to be made. Part 3 outlines an extensive series of fundamental principles that must be understood in order to derive the maximum benefit from a quality assurance program. And finally, Part 4 presents a plan of action that, if conscientiously followed, will significantly increase the effectiveness of transportation quality assurance practices nationwide.

PART 1: FACING THE PROBLEM

Few would argue with the statement that U.S. roads and bridges are among the nation's most valuable assets. In a recent column (*1*), George Will states that transportation (all types) makes up 18 percent of the gross national product, employs one-tenth of the work force, and accounts for 15 to 30 percent of the cost of agricultural products. In another publication, (*2*, p. 1-4), it is stated that the total replacement value of the nation's roads and bridges is estimated to be between $1 and $3 trillion and that any measure that could improve their performance and durability by even 1 percent would result in savings of billions of dollars. Although this latter reference was focused on the projected benefits of research, there is another means by which the performance and durability of roads and bridges can be dramatically improved. The research is complete, and the results—well documented by actual data from many states—suggest that the expected improvement is substantially greater than 1 percent. We are referring to statistical quality assurance (SQA), a body of knowledge and procedures that provides a strong incentive to industry to use state-of-the-art techniques to obtain high-quality construction.

FACING ISSUES

In a recent paper on highway safety, Hauer (*3*, pp. 241–267) makes some telling points about the tremendous impact of the transportation engineering profession on the overall performance of the roads, bridges, and other modes of transportation under its purview. The author stresses the importance of facing up to certain responsibilities and the consequences of not facing up to these responsibilities. We shall take a similar approach to express our own concerns about what we perceive as serious shortcomings in the area of quality assurance. We shall then outline a series of facts that must be recognized and responsibilities that must be faced in order to derive the full benefits that quality assurance has to offer.

GOALS OF QUALITY ASSURANCE

In the field of transportation, the term "quality assurance" is generally associated with a comprehensive program to achieve conformance with established desired quality levels for design and construction. This program involves people, materials, equipment, procedures, and the optimal use of these resources. Probabilistic concepts and statistical acceptance procedures are frequently used. The following scope statement for the Transportation Research Board Committee A2F03 states the goals quite clearly:

> This committee will concern itself with all aspects of total quality management in the transportation field. It will endeavor to foster responsible leadership in the application of both engineering and statistical knowledge toward sound, practical, and effective quality assurance procedures. It will develop and promote methods to achieve high quality design, construction, and maintenance at the lowest possible overall cost. These efforts will include, but not necessarily be limited to, end-result and performance-related specifications, statistical quality assurance and control techniques, acceptance sampling, accuracy and precision of tests, optimal use of limited resources, cost-effectiveness of quality assurance procedures, evaluation of consensus standards, preparation of monographs, and educational programs related to these topics.

This is an admirable set of goals, but if they are to be realized, we believe that a change in thinking and priorities

K. C. Afferton and R. M. Weed, New Jersey Department of Transportation, 1035 Parkway Avenue, Trenton, N.J. 08625. J. Freidenrich, The RBA Group, 1 Evergreen Place, Morristown, N.J. 07962-1927.

will be required. The old attitude of "close enough for highway work" will not do. This not only downplays the importance of high-quality construction, but also serves to discourage the use of modern, effective statistical procedures. If anything, there needs to be a greater willingness to seek out modern technology and use it to its fullest advantage.

WORK ETHIC AND PROFESSIONAL ETHICS

Much has been written recently about the trade deficit and the invasion of American corporations and markets by foreign interests. Various explanations have been offered, but it comes down to this: the United States has been outdone at what used to be its strength—developing and applying modern technology.

Many have observed what has been described as a persistent and progressive deterioration of the work ethic. A rather lengthy list of examples could be compiled, beginning with a general decline of educational standards and ending with an erosion of fundamental values that has led to a whole host of corporate and political ills. The problem is real, it is pervasive, it is present in both the public and private sectors, and it is unlikely to improve unless specific action is taken to correct it.

In the field of quality assurance, the problem manifests itself in a particularly troublesome way. In a discipline dedicated to the pursuit of excellence, it seems totally inappropriate to tolerate specifications and consensus standards that are far from excellent. If demands for excellence are to be made of the construction industry, it is imperative that engineers be willing to demand the same of themselves in the development of the specifications and standards that govern this work.

Ironically, all the necessary statistical tools are well developed and readily available. What is lacking, however, is a widespread willingness to use them. To take advantage of the benefits that modern methods have to offer, engineers can no longer afford to cling to the old ways of doing things just because these ways are familiar and comfortable. Those unfamiliar with the mathematical principles underlying SQA procedures may find it difficult to realize just how inadequate many current practices are.

Leaders within the transportation field must invite an open and thorough scrutiny of current practices and must insist that improvements be made where necessary. To do anything less would be a breach of both professional ethics and public trust.

NEW JERSEY'S EXPERIENCE

Our own experience with SQA, obtained over a period of approximately 20 years, is the basis for what we advocate. We present a brief history here with the belief that it will be both informative and helpful.

A congressional investigation (*4*, p. 3) in 1962 uncovered many cases of nonconformance in highway construction throughout the country at about the same time the (then) AASHO Road Test provided a wealth of statistical data relating quality measures to performance (*5*). This led to the realization that various statistical measures can effectively describe the characteristics that are desired and provided the motivation to commence the development of statistically based specifications.

New Jersey was one of several states that began to explore the benefits of this new approach. State design engineers were quick to recognize that, for many highway items, it simply was not possible to define a single level of quality that clearly separated acceptable and unacceptable work. Furthermore, it usually was not practical to require removal and replacement of an item that was only marginally deficient, but on the other hand, neither was it appropriate to accept such an item and pay full price for it. Statistical specifications provided a convenient and practical way to accept these items for a prearranged reduced level of payment.

Random sampling plans, statistical acceptance procedures, and adjusted pay schedules were first developed for various properties of bituminous concrete. As time went on, similar procedures were developed for pavement thickness and surface smoothness and for portland cement concrete (PCC) strength.

The concrete specification was the first in New Jersey to incorporate a bonus provision, paying up to a maximum of 102 percent for exceptionally high strength. The most recent New Jersey statistical specification includes a combined acceptance procedure for thickness, strength, and surface smoothness of rigid pavement. Surface smoothness dominates and, provided all three parameters are under control, a maximum pay factor of 103 percent is obtainable. A unique feature of this specification is that, within reasonable limits, excesses and deficiencies in thickness and strength are allowed to compensate for each other.

SQA has worked well for the New Jersey Department of Transportation (NJDOT), which continues to use it and strongly endorses it. NJDOT administrators have also learned what to expect when an agency embarks on such a program. First, there is the usual resistance to change, sometimes from within the transportation agency but usually from the various industry organizations outside. The approach is new to many people and it is normal to fear the unknown. Next comes the learning process, in which both the transportation agency and the construction industry become familiar with the new procedures. There are some growing pains—nothing worthwhile comes without effort—but this phase usually proceeds more smoothly if the transportation agency and the construction industry engage in cooperative field trials. And finally comes the actual implementation. On the basis of NJDOT's experience, bid prices may go up a little but quality usually increases dramatically. As time goes on, bid prices usually tend to stabilize near their former levels as contractors become more familiar with the new specifications.

CASE STUDY

The PCC specification, the most recent of NJDOT's statistical specifications to be widely implemented, serves as a case study for a number of fundamental concepts that are described in detail in Part 3 of this paper. Although both slump and air entrainment are routinely tested and used to accept or reject the material at the job site, final acceptance is based on compressive strength. The percent defective (the percentage of the concrete estimated to have strength below the specifica-

tion limit) was chosen as the control parameter because it is believed to be closely related to performance. Historical data were reviewed to determine the levels of percent defective that have been associated with both satisfactory and unsatisfactory performance. It was decided to raise existing quality levels slightly, so new specification limits were set accordingly. Any quality-performance data that could be found (or estimated) were used along with engineering economics principles to develop a pay schedule that is rational, defensible, and, in our opinion, fair and equitable. This includes a bonus provision to provide still additional incentive to achieve high-quality construction. An acceptance plan based on standard variables procedures is used, and the operating-characteristic curves were checked to confirm that the specification would perform as intended. These developments are described in greater detail in a separate publication (*6*).

The development of this specification produced several new ideas, such as the bonus provision, that were a distinct departure from some earlier NJDOT practices. A variety of difficult decisions had to be made, and they were made much easier by the thorough technical approach that was taken. Computer simulation, in particular, permitted management to see exactly what the capabilities of the specification were and the probable consequences of various provisions that were under consideration. This put management in a much stronger position to provide the leadership necessary to overcome the initial skepticism that often accompanies an effort such as this.

Although developing a technically sound acceptance procedure is an essential first step, several other steps are equally important. If cooperation is to be expected from those who will ultimately be responsible for enforcing the requirements of the specification, management must clearly communicate what is wanted and why it is important to the organization. Toward this end, several training sessions were held by NJDOT to explain how and why statistical specifications work and what they are expected to accomplish. Construction industry representatives were frequently invited to these sessions to give them advance notice of the plans and to give them the opportunity to obtain additional information and provide input. Before the specification was implemented, several field trials were conducted to familiarize everyone with the new procedures. Comments and suggestions were solicited, both from within NJDOT and from the construction industry. This produced several useful insights and seemed to foster a greater sense of involvement and cooperation among all concerned. The final implementation was accomplished in a gradual fashion by reducing all pay adjustments by half for the first two projects. Ultimately, we believe that this broad involvement and methodical approach contributed to a generally smooth implementation when the specification was finally adopted.

For the first project, there was approximately 10,000 yd^3 of concrete, consisting of both pavement and structures, with an in-place value of about $3 million. The successful bidder subcontracted the concrete production to an independent transit mix producer. Several different mix designs were prepared and tested before construction to demonstrate the producer's ability to meet the specification requirements. Because of the careful preparation and a conscientious effort by everyone involved, the quality remained consistently high throughout the project. The compressive strength averaged approximately 1,000 psi higher than would have been necessary to receive 100 percent payment, and as a result the contractor earned bonus payments totaling approximately $30,000.

The second project, which was of approximately the same size and type, went to a different contractor and producer. The performance in this case was still better, and this contractor also received essentially the maximum amount of bonus payment.

NJDOT was well satisfied with these results because it was believed that more than comparable value was received in terms of extended service lives of the pavement and structures built under this specification. It is presumed that the contractors and producers were also pleased, both with the monetary value of the bonus and with the recognition for having run such well-controlled projects.

The relationship between the contractors and producers was proprietary, and it is not known what arrangements were made to share either bonuses or any potential pay reductions. It is believed that the producers charged more for concrete supplied under these contracts, but this information, too, was proprietary. The contractors' unit prices were not abnormally high, however, and the bids for the projects as a whole were below the anticipated cost for the work.

After the successful completion of these two projects, some minor improvements were made and the specification was then adopted for all future work. After a suitable amount of field experience has been obtained, it is planned to again seek feedback to determine if any further refinements are desirable.

STATUS OF SQA NATIONWIDE

Where does SQA stand nationwide? On the basis of the results of a 1984 survey (R.M. Weed, unpublished data), approximately one-half of the states are actively using this approach and another one-fourth have statistical specifications in various stages of development. Also, virtually every state that has tried SQA continues to use it. Clearly, many consider this approach to be preferable to the earlier "method" specifications under which transportation agencies provided detailed instructions, supervised the construction operations closely, and bore most of the responsibility for the outcome. And no wonder—statistical specifications are easier to write (just describe the desired end result in statistical terms), easier to interpret (no vague terms like "reasonably close conformance"), easier to enforce (clear separation of responsibilities for control and acceptance), and easier to apply (pay adjustment for defective work is predetermined; no negotiations are required). An additional benefit of SQA is the data it produces. Whereas historical data collected in conjunction with method specifications have been notoriously unreliable, SQA specifications produce accurate data obtained with valid random sampling procedures. This is particularly important if these data will be analyzed at a later date to develop better performance-related specifications.

In spite of this progress, with which the transportation profession can justifiably be proud, much remains to be done. Although SQA appears to be performing well, there is a distinct possibility that this might be illusory in many cases. Many current practices and published standards pertaining to

SQA are far from optimal, and some may actually be incorrect (*7–11*). Unfortunately, there often is no immediately obvious indication that a statistical procedure is being misapplied. Instead, there may simply be a false sense of security, which most likely will be paid for in terms of premature failures and costly repairs.

BASIC FACTS AND RESPONSIBILITIES

How, then, can states be reasonably assured of deriving the maximum benefits that SQA has to offer? It seems to us that engineers must learn to face up to several facts and be willing to accept certain responsibilities. The facts are these:

1. SQA is one of the most useful tools that a transportation agency has at its disposal. Provided that the engineering designs are adequate, it can virtually guarantee the performance and durability of the items to which it is applied. This is usually accomplished by encouraging high-quality construction initially, but also by recouping the anticipated future repair costs in the form of pay reductions when quality is substandard.
2. The quality of SQA specifications has a direct bearing on the quality of construction that is produced.
3. Any failure to derive the full benefit of SQA is not a failure of the techniques themselves, but rather a consequence of an inappropriate application. Shortcuts taken by practitioners unfamiliar with statistical methods can produce procedures that are inefficient at best or ineffective at worst. Today, many current practices are far from optimal, falling well short of their potential in terms of simplicity, effectiveness, or economy.
4. The upgrading of existing applications, combined with an increase in the number of effective applications, would enhance both the durability and performance of the items to which they are applied and would produce substantial economic benefits.
5. SQA techniques can be presented in a simple and straightforward manner. They need not be complex, nor is an extensive statistical background necessary to use them correctly.
6. Statistical principles are universally applicable. To take advantage of them requires only that the basic underlying assumptions be reasonably satisfied and that the prescribed steps of the procedures be followed correctly. This is relatively easy to accomplish for most SQA applications.

These are the responsibilities:

1. Leaders in the transportation field must take the initiative to thoroughly investigate any measure that offers so great a potential return as the proper use of SQA.
2. If transportation leaders decide to endorse such an approach, whether it be the implementation of a new program or merely the enhancement of an existing program, their support should be visible and active. An effective SQA program is truly a "mind-set" that requires a commitment to excellence by every member of an organization from the top down.
3. The successful application of any technical discipline requires a working knowledge of that discipline. Just as bridge designers are expected to have a thorough knowledge of structural analysis, the developers of SQA specifications must have a thorough grounding in statistical principles. It is up to management to insist that appropriate personnel acquire and use the necessary skills. Because NJDOT recognized how important a command of this subject was to the development of a sound and defensible program, a new set of job specifications was created—the Statistical Engineer Series—to ensure that NJDOT would have staff with the necessary combination of engineering and statistical knowledge. It is this group that handles the technical aspects whenever a new SQA specification is developed.
4. An organization that is committed to quality assurance has an obligation to become familiar with the state of the art and to apply it in a thoroughly professional manner. SQA is now a totally scientific activity; there is no longer a need for guesswork or trial-and-error approaches that may, in fact, fail to provide the level of quality assurance that the agency expects.
5. Standards-writing groups, such as AASHTO and ASTM, must be especially diligent. Anything that is published as a procedural guide must be held to the highest of professional standards.
6. Although the users of SQA methods do not need to have as extensive a technical background as the developers of the procedures, they are much more likely to be willing and enthusiastic supporters of these methods if they have a basic understanding of how and why they work. In our opinion, it is essential that basic educational programs be provided for all who will work with SQA.
7. To be truly successful, a quality assurance program requires the cooperation of all parties involved, including the construction industry, which must be treated in a fair-minded fashion. 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 new procedures. The transportation agency must be continually critical of its own performance, must be prepared to acknowledge mistakes, and must act quickly to correct any discovered deficiencies to guarantee the program's continued success.

NEED FOR A NATIONAL POLICY

If transportation leaders can be made aware of these facts, and a sufficient number are willing to accept the attendant responsibilities, this should provide the motivation necessary to establish a national policy on quality assurance. The potential benefits are many. A consistent approach nationwide, using the best methods in the most effective ways, would produce a general increase in construction quality and result in substantial cost savings and other long-term benefits. Consistency from state to state would tend to reduce confusion, both on the part of the engineers who write statistical specifications and the contractors whose work is governed by them. This could lead to simpler specifications, lower bid prices, and smoother implementation. It would almost certainly im-

prove contractual relations and lessen the resistance typically encountered from the construction industry. Uniformity from state to state may also aid in the establishment of a national data base that might ultimately provide the information necessary to develop better performance-related specifications. A uniform approach known to be scientifically sound may also tend to increase the credibility of those agencies that use it. We believe that this would lead to a heightened respectability for statistical methods in general and might provide the incentive needed to encourage the remaining 25 percent of the states to discover the advantages of SQA.

VISION AND DIRECTION NEEDED

If such a national policy is to be established, we believe that there will have to be distinct changes in the organizational culture of many transportation agencies. Of the various elements essential to an effective program, leadership is perhaps the most important. It is the leadership that first envisions the goals and then provides the direction to achieve them. As discussed in a recent article (*12*), the organizational culture ultimately reflects the leaders' values, their sense of what conditions ought to exist, and their readiness to work toward these goals. Consequently, leaders must have a clear vision of exactly what is to be accomplished in order to create the type of environment most conducive to the achievement of the goals of the organization.

In the field of quality assurance, if realistic and effective goals are to be established, it is absolutely essential that certain basic requirements be met first. Either the leaders themselves must have a solid technical grasp of both the capabilities and limitations of SQA, or they must acquire qualified staff or consulting assistance from outside their respective organizations, or they must obtain suitable formal training for themselves or their staff. Ideally, the individuals who either provide or acquire this expertise should also have a background in engineering so that they will have a practical understanding of the construction work to which the quality assurance techniques will be applied. The potential benefits of a sound quality assurance program are of such magnitude that it warrants nothing less than the same high level of expertise devoted to other engineering and scientific disciplines.

And finally, if efforts in this direction are to be successful, a sense of responsibility and accountability must be established. It is up to leaders who undertake this task to set up some sort of technical review process to provide periodic feedback to assure that their goals are in fact being accomplished.

SUMMARY AND PREVIEW

The highway system represents an investment valued in trillions of dollars. SQA, widely used to govern both construction and maintenance work, has considerable unrealized potential because of a variety of less-than-optimal applications. A general upgrading of current practice coupled with a wider use of effective procedures might produce long-term benefits measured in billions of dollars.

Leaders at all levels—heads of transportation departments, administrators of federal programs, chairmen of Transportation Research Board committees and technical standards groups, managers of research programs—have an obligation to thoroughly consider the benefits that might be derived from a more rigorous application of modern quality assurance methods. In our opinion, this could best be accomplished through the creation of a national task force to provide the necessary vision, direction, and leadership. Individuals with the required combination of multidisciplinary skills could be recruited on a voluntary basis from the transportation, academic, and industrial communities.

To aid those who wish to explore this proposal further, Part 2 of this paper warns of several obstacles that might impede either the effective application of SQA procedures by individual agencies or the development of a sound national policy on quality assurance. It also presents several examples of shortcomings in existing standards and specifications and makes recommendations to improve them. Part 3 presents a series of fundamental principles that must be understood and applied correctly if truly effective quality assurance programs are to be established. Part 4 then draws from the foregoing material to develop a plan of action to assure the effective application of these concepts.

REFERENCES (Part 1)

1. G. Will. Is Road-Building Still a GOP Tradition? *Trenton Times*, March 12, 1990, Section A, p. 8.
2. *Innovation—A Strategy for Research, Development, and Technology Transfer*. American Association of State Highway and Transportation Officials, Final Draft. Washington, D.C., Aug. 1989.
3. E. Hauer. A Case for Science-Based Road Safety Design and Management. In *Highway Safety: At the Crossroads* (R. E. Stammer, ed.), American Society of Civil Engineers, New York, 1988.
4. *NCHRP Synthesis of Highway Practice 38: Statistically Oriented End-Result Specifications*. TRB, National Research Council, Washington, D.C., 1976.
5. *Special Report 61G: The AASHO Road Test, Report 7, Summary Report*. Highway Research Board, National Research Council, Washington, D.C., 1962.
6. R. M. Weed. *Statistical Specification Development*, 2nd ed. HPR Study 7788. New Jersey Department of Transportation, Trenton, 1989.
7. R. M. Weed. Analysis of Multiple Acceptance Criteria. In *Transportation Research Record 745*, TRB, National Research Council, Washington, D.C., 1980, pp. 20–23.
8. R. M. Weed. Analysis and Application of Correlated Compound Probabilities. In *Transportation Research Record 792*, TRB, National Research Council, Washington, D.C., 1981, pp. 41–44.
9. R. M. Weed. Adjusted Pay Schedules: New Concepts and Provisions. In *Transportation Research Record 986*, TRB, National Research Council, Washington, D.C., 1984, pp. 32–38.
10. P. A. Kopac. Current Practices in Acceptance of Bituminous Concrete Compaction. In *Transportation Research Record 986*, TRB, National Research Council, Washington, D.C., 1984, pp. 43-46.
11. R. M. Weed. Revision of a Flawed Acceptance Standard. In *Transportation Research Record 1056*, TRB, National Research Council, Washington, D.C., 1986, pp. 21–35.
12. P. DeBaylo. Organizational Culture and Leadership. *Quality Focus* (newsletter of Princeton Section of American Society for Quality Control), Winter, 1990, pp. 1 and 3.

PART 2: OBSTACLES TO OVERCOME

Many obstacles stand in the way of a successful quality assurance program. They may be technical, managerial, behavioral, political, cultural, budgetary, regulatory, or any combination of these.

We believe that the technical problems are relatively minor. The relationships between design values and ultimate performance are reasonably well understood. Test methods for engineering materials and properties are generally well defined. Optimal statistical techniques for sampling a product and estimating its quality are well established and easy to apply. Although advances in these areas may still be made, all the necessary technical tools are now in place to support an effective quality assurance program. The only significant problem in the technical category is the existence of inadequate standards and guide specifications (*1*) that could mislead those unfamiliar with statistical acceptance sampling procedures.

The fact that optimal SQA techniques are not being widely used is one of several problems that fall into the managerial category. Transportation leaders must become aware that better tools exist and that a very real price—in terms of wasted sampling effort, greater risk of accepting poor quality, and some equally serious but less obvious consequences—is being paid when inferior methods are used. Perhaps the underlying problem is that there has been a lack of visible leadership that would serve to encourage and motivate the development of a comprehensive and sound policy on transportation quality assurance.

Whenever a quality assurance program is implemented, resistance will inevitably be encountered. Sometimes it will be internal, from employees who are more comfortable with the older ways of doing things, but more often it will be external, from contractors who are apprehensive about any system that can affect their profitability. Industry lobbies can be vocal and influential, and it is essential that transportation agencies have a sound plan of action before undertaking such a program. Specific guidance on these matters is provided in Parts 3 and 4 of this paper.

Political factors also play a role. Public perception can sometimes be a stronger motivating force than sound scientific rationale. Unfortunately, decisions that make the greatest technical sense in the long term are sometimes unpopular in the short term.

"Made in the USA" used to be the hallmark of quality. But a decline in academic achievement (particularly in science and mathematics), a work ethic that seems to be predicated on being just good enough to get by, a management that is technically indifferent, and a willingness to tolerate mediocrity now pose major challenges to the United States' reputation as a technological leader. A similar situation is evident in the manner in which SQA has been applied in the transportation field.

Budgetary problems have become increasingly acute in recent years, resulting in both personnel and equipment shortages for many transportation agencies. Although managers have little control over the amount of funding available, they do have fairly direct control over how the funds are spent. Total quality management (TQM) concepts in general, and SQA procedures in particular, can help ensure that these funds are used resourcefully.

And finally, laws and regulations can sometimes limit potentially desirable applications. For example, FHWA supports positive incentive (bonus payment) clauses for superior quality, and several states have found this to be a practical and effective approach (*2*). However, such clauses apparently are not permitted by the laws of some states. Similarly, as discussed in the report of the European Asphalt Study Tour given at the 1991 TRB Annual Meeting, some of the contractual relationships that have been demonstrated to work well in Europe almost certainly would not comply with current laws in the United States. Although these may be the hardest obstacles to overcome, they need not be considered insurmountable. If and when there is ample evidence that laws need to be changed to keep pace with modern technology, government leaders should be able to bring their collective influence to bear on such issues.

Of the several types of obstacles just discussed, the majority by far are well within the control of a management that is committed to excellence in quality assurance. In the remainder of Part 2, we will explore several of these obstacles in detail and the steps that can be taken to overcome them. Specific topics to be addressed are the following:

- Little demand for excellence,
- Complacency about existing SQA practices,
- Uncertainty about effectiveness of SQA methods,
- Inadequate procedures and practices,
- Poor teaching of statistical methods,
- Resistance from within the transportation agency,
- Opposition from the construction industry,
- Political factors,
- Work ethic and cultural attitudes, and
- Conveying the wrong messages.

LITTLE DEMAND FOR EXCELLENCE

Occasionally, a situation arises that is sufficiently disturbing that conscientious people feel compelled to do something about it. The status of highway quality assurance is one of those situations. When a body of knowledge exists that is both extremely useful and easy to use and that body of knowledge is either widely misused or not used at all, then something is basically wrong. When national standards-writing committees demonstrate neither a concern for these problems nor a willingness to resolve them, something needs to be done.

Although statistical acceptance procedures are widely used—about three-fourths of the states either actively use them or have them in various stages of development—there is great disparity in the manner in which they are applied (*3*). Many are far from optimal, some are ineffective, and a few are blatantly incorrect (*1,4*).

It is not much of a stretch of the imagination to see a parallel to the American manufacturing industry of the 1960s. Writers such as Deming (*5*) and Juran (*6*) tried to tell industry leaders that they could dramatically improve their operations. But

business was good back then; American management didn't have to be particularly effective to capture a large share of both the national and international markets. They were—as Malcolm Baldrige characterized them—"fat, dumb, and happy" (7). So Deming and Juran took their advice to Japan, with incalculable cost to the American economy.

Is the highway profession, which is currently performing well below its capability in the area of quality assurance, poised for a similar decline? No one knows, of course, but there certainly are some ominous signs. In 1988, the U.S. Department of Transportation classified only 57 percent of the Interstate highway system in good condition, and the rest ranged from "fair" to "wretched" (*8*). The non-Interstate portions of the highway system tend to be in somewhat poorer condition and the situation with highway bridges is similarly alarming (*9*).

Given the status of quality assurance practices in the United States, it is not unreasonable to lay some of the blame for this state of disrepair on the failure to provide sufficient incentive to produce high quality at the time of construction. (For example, if a statistical acceptance procedure for rigid pavement encourages an increase in as-built thickness of ¼ in. rather than a deficiency of the same amount, this will increase the load-carrying capacity of the pavement by about 35 percent, affecting service life accordingly.) Although nothing can be done about the events of the past, they can still provide useful guidance to encourage better practices in the future. What are needed now, particularly during this time of shrinking resources, are the leadership and resolve to insist that modern quality assurance technology be used to its fullest advantage.

COMPLACENCY ABOUT EXISTING SQA PRACTICES

"If it ain't broke, don't fix it!" This frequently heard expression is generally interpreted to mean that a system that appears to be working well should not be disturbed. This advice seems sound enough and there undoubtedly are situations in which it would be wise to heed it. However, a number of recent writers on quality management, most notably Deming, caution that there are times when this advice may be especially inappropriate (*10*).

Perhaps one of the most difficult obstacles to overcome is a general lack of awareness of the inadequacy of many current SQA standards and specifications. Unfortunately, neither the infrequent occurrence of reject lots nor the absence of short-term failures is a guarantee that an existing SQA program is performing properly. Unless the operating characteristic curves (described in Part 3) have been constructed, there usually is no way to be sure that a statistical acceptance procedure is sufficiently capable of distinguishing between satisfactory and unsatisfactory work. For example, if the acceptance procedure happens to be weak (or even if it is strong but the specification limits are not suitably restrictive), moderately defective work would not be detected and the project would have the appearance of being well controlled. When this happens, pavements may begin to fail after 10 years instead of 20, or structures may begin to crumble after 25 years instead of 50, and the cost of such premature failure can be substantial. By the time the failure occurs, however, the cause will be difficult to ascertain. Very probably, it will be further obscured by the additional routine maintenance that these items will receive throughout their lifetimes.

To guard against this potential weakness, it is necessary to examine all statistical acceptance procedures critically, even those that appear to be working well. It is relatively easy to construct the operating characteristic curves and these, at least, give the transportation agency the information needed to decide whether the acceptance procedures are providing the desired degree of protection.

UNCERTAINTY ABOUT EFFECTIVENESS OF SQA METHODS

An oversight, for which proponents of SQA are not entirely to blame, is the failure to establish definitively that quality assurance programs are cost-effective. The type of controlled studies that would be capable of demonstrating this are extremely difficult to design and, to our knowledge, have never been attempted.

There are, however, other measures by which to judge the effectiveness of a quality assurance program. A primary one is its effect on the general quality level of the items to which it is applied. A second but equally important consideration is the cost of implementing and maintaining such a program. A third measure is the effort that administrators are willing to make to actively support SQA methods.

It has been our experience that the introduction of a new quality assurance specification usually produces a dramatic improvement in quality. As described in the case study in Part 1 of this paper, there was a general increase of between 1,000 and 1,500 psi in concrete strength levels when the specification was implemented with the pay adjustment clause in effect. Judging from responses obtained from a 1989 survey of many state transportation agencies (O. Riley, unpublished data), this result is fairly universal.

The more difficult question to answer is whether the expected increase in quality justifies the cost of achieving it. In the absence of the necessary data, we can only offer an opinion—one that seems to be shared by our counterparts in the approximately three-fourths of the states that are actively pursuing SQA methods—that the advantages far outweigh any possible disadvantages. As noted in Part 1 of this paper, statistical specifications are easier to write (just describe the desired end result in statistical terms), easier to interpret (no vague terms like "reasonably close conformance"), easier to enforce (clear separation of responsibilities for control and acceptance), and easier to apply (pay adjustment for defective work is predetermined; no negotiations are required). An additional benefit that comes almost as a bonus is that the existence of a formal SQA program ensures the development of valid data bases that may be useful for the development of better performance models in the future.

Finally, in regard to the amount of effort that transportation administrators are willing to expend to overcome the skepticism and resistance that inevitably accompany a new program of this type, we offer the following empirical evidence. In spite of outside resistance that is frequently well organized,

and internal resistance that is often quite vocal, the use of SQA methods has been steadily increasing over a period of approximately 20 years. Also, virtually every agency that has tried them continues to use them. We believe that these facts, because they reflect the collective opinion of many leaders throughout the transportation field, provide a particularly strong endorsement of the effectiveness of SQA methods.

INADEQUATE PROCEDURES AND PRACTICES

For a variety of reasons, many current quality assurance practices are far from what they should be. When inferior methods are used, the transportation agency—and ultimately the motoring public—pays for this in one way or another:

1. Sampling rates may be larger than necessary for the degree of protection afforded, wasting personnel, materials, and storage space.
2. The risk of accepting poor quality may be significantly greater for a given sampling rate, increasing the likelihood of premature failure of various construction items.
3. Inferior methods occasionally produce inconsistent results and sometimes do not provide the proper incentive and reward to the conscientious contractor (*2*). This will eventually become apparent, even to those unfamiliar with statistical methods, and is responsible to a large degree for breeding a general distrust of SQA programs.
4. This general distrust tends to produce an adversarial relationship between transportation agencies and the construction industry rather than the cooperative climate that is desired. This can escalate to potential litigation if contractors sense that the acceptance procedures are not technically sound and defensible.
5. Less-than-optimal methods, if they are unduly severe, will produce such strong resistance from the construction industry that political representatives may be called upon to intervene. This may also lead to an undesired inflation of bid prices.
6. In general, methods lacking a clear scientific rationale are undesirable for several reasons. They are difficult to justify, difficult to explain, and difficult to defend. They tend to be confusing to all concerned. This makes it all the more difficult to build confidence in the quality assurance program and generate the support and cooperation needed within the transportation agency itself.

The examples that follow are but a few of the literally dozens that could be included. Although no single example by itself is damning, the fact that so many examples exist is indicative of the need for major reforms.

Example 1: Wasted Sampling Effort

When SQA specifications were first developed in the 1960s, most transportation agencies chose to use the range rather than the standard deviation as the measure of variability. The range was easier to understand, and hand calculators that could compute the standard deviation with a single keystroke were not yet available. Another simplification was the occasional use of attributes procedures, which avoided the need for any statistical calculations and involved only the counting of the number of failing tests. The different types of acceptance procedures are discussed in more detail in Part 3. The important thing to note here is that there is a considerable difference in the efficiency of these methods. They can all be designed to be effective—to distinguish between satisfactory and unsatisfactory work at whatever level of risk is believed to be appropriate—but to be equally effective, they require different sample sizes.

Table 1 is a comparison of these three approaches—a plan based on the standard deviation, one based on the range, and an attributes plan—all designed to control percent defective (the percentage of the lot falling outside specification limits). The effectiveness of these plans can be judged by their operating characteristic curves, the data for which are shown in Table 1. It will be noted that the probabilities of acceptance for the three plans at all levels of quality are virtually identical; that is, the plans are all equally effective. The measure of efficiency is provided by the sample sizes given in the column headings ($N = 10$, $N = 12$, and $N = 15$). The range plan is substantially less efficient than the standard deviation plan because it requires a 20 percent larger sample to provide the same degree of protection. The attributes plan is considerably less efficient, requiring a 50 percent larger sample.

At a time when more than half the states are experiencing severe budget shortfalls (*11*), any plan that does not make the most efficient use of the available data must be regarded as inadequate. There are many existing range plans, and possibly a few attributes plans, that could be readily converted into more efficient standard deviation plans, thus affording an appreciable savings in personnel, materials, and storage.

Example 2: Potentially Ineffective Plans

Although the three acceptance plans just discussed do account for variability of the product, many currently used plans do not (*4*). These plans fall into two general categories:

1. Plans that are based only on the average of the test values and
2. Standard deviation plans that assume a constant, known value for the standard deviation.

TABLE 1 COMPARISON OF EFFICIENCY OF THREE ACCEPTANCE PLANS

	Probability of Acceptance		
	Variables Plans		
Percent Defective	Standard Deviation Method ($N = 10$)	Range Method ($N = 12$)	Attributes Plan ($N = 15$)
10	0.99	0.99	0.99
20	0.82	0.82	0.84
30	0.50	0.50	0.52
40	0.21	0.21	0.22
50	0.06	0.06	0.06
60	0.01	0.01	0.01
70	0.0	0.0	0.0

These plans may not provide the quality that is desired. Many engineers in the transportation field believe that the ultimate performance of the final product is influenced by both the average and the variability of the construction process. The underlying rationale is discussed further in Part 3 of this paper and also in American Concrete Institute (ACI) Standard 214 (*12*).

Example 3: Inadequate Technical Aids

If SQA is to be promoted and widely adopted, it is up to its proponents to make it as user friendly as correct techniques will allow. Proper training can make SQA easy to understand, but a variety of technical aids is necessary to make it easy to use. As an example of a technical aid that is not properly designed, consider Table 4A in Method B of AASHTO Standard R9 (*13*). The heading of the table indicates that its purpose is to provide an estimate of the percentage of a lot falling within specification limits as a function of the quality index (Q-statistic). A portion of this table is reproduced here as Table 2.

What is obvious upon inspection is that this table is constructed in a backward manner. Whereas it is customary to construct such tables with the input variable around the perimeter and the output variable in the body of the table, exactly the reverse has been done here. Another major drawback is that not all potentially useful sample sizes have been included. To see how inconvenient this is for the user, just try determining the percent within limits for $N = 5$, $Q = 1.46$ or $N = 6$, $Q = 1.53$. Both cases require interpolation, the latter one in two directions.

As an example of how such a table should be constructed, consider Table 3. Again try the same two examples to see how much more user friendly this version of the table is. Examples such as this stress the importance of having technical aids designed by individuals familiar with both theory and practice. It also is easy to understand how the users of the improved version of the table might end up being much more supportive of SQA methods.

Another very useful form for this table, generated by an appropriate computer algorithm, is presented as Figure 1. This particular version is based on percent defective (the complement of percent within limits). This is an extremely compact form of the table for a single sample size (useful in specification documents) in which the Q-values are listed in increments of 0.1 in the left-hand column and 0.01 across the top. A more extensive series of new tables is in preparation (*14*), and additional examples are provided in Part 3.

TABLE 2 POORLY CONSTRUCTED TABLE FOR ESTIMATING PERCENT WITHIN LIMITS

Percent Within Limits	Q-Values				
	$N = 4$	$N = 5$	$N = 7$	$N = 10$	$N = 15$
...	...	...	...	...	...
97	...	1.55	1.67	1.74	...
96	...	1.49	1.59	1.64	...
95	...	1.45	1.52	1.56	...
94	...	1.40	1.46	1.49	...
93	...	1.36	1.40	1.43	...
...	...	...	...	...	...

TABLE 3 PROPER FORM FOR TABLE FOR ESTIMATING PERCENT WITHIN LIMITS

Q-Value	Percent Within Limits				
	$N = 4$	$N = 5$	$N = 6$	$N = 7$	$N = 8$
...	...	...	...	...	...
1.46	...	95.40	94.50	94.07	...
1.47	...	95.61	94.67	94.23	...
1.48	...	95.81	94.85	94.40	...
1.49	...	96.01	95.02	94.56	...
1.50	...	96.20	95.19	94.72	...
1.51	...	96.39	95.36	94.87	...
1.52	...	96.58	95.53	95.03	...
1.53	...	96.77	95.69	95.18	...
...	...	...	...	...	...

Example 4: An Inadequate Standard

Some further comments are warranted in regard to Method B of AASHTO Standard R9. This work dates back to the 1960s when the early development of SQA specifications was based more on trial and error than on an actual understanding of statistical principles. About the most diplomatic thing that can be said about Method B is that it simply is not mathematically sound (*1*; personal communications: C. Antle, Pennsylvania State University; P. Irick, consultant; E. Schilling, Rochester Institute of Technology; O. Pendleton, Texas Transportation Institute; W. Strawderman, Rutgers University). It is the hybrid of two methods (averages and variables) that should not be combined and it misuses both. If the quality assurance profession is to command the respect necessary to be widely supported, it must be uncompromising in the demands it places on itself for scientific integrity. If the transportation profession is to demand excellence of the construction industry through the use of SQA procedures, it must be no less demanding of itself in making these procedures clear, effective, and technically sound. Method B fails to meet this standard of excellence and, in our opinion, it is not salvageable as a valid procedure.

Example 5: Equitable and Defensible Pay Schedules

Adjusted pay schedules have been a source of controversy since they were first introduced in the 1960s. On the one hand, they provide the most practical way for transportation agencies to deal with marginal quality, which, like it or not, will occur from time to time. It usually makes more sense to accept a slightly deficient construction item for some fraction of the full price than it does to require that it be torn out and replaced. Adjusted pay schedules also provide an effective way to encourage and reward superior quality with bonus clauses. On the other hand, they have often been controversial because no clearcut scientific rationale has been universally accepted for determining the appropriate amount of pay adjustment. This is evident from the great disparity among pay schedules used by different transportation agencies across the country (*3*).

VARIABILITY-UNKNOWN PROCEDURE — SAMPLE SIZE 10 — STANDARD DEVIATION METHOD

Q	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	50.00	49.62	49.23	48.85	48.46	48.08	47.70	47.31	46.93	46.54
0.1	46.16	45.78	45.40	45.01	44.63	44.25	43.87	43.49	43.11	42.73
0.2	42.35	41.97	41.60	41.22	40.84	40.47	40.09	39.72	39.34	38.97
0.3	38.60	38.23	37.86	37.49	37.12	36.75	36.38	36.02	35.65	35.29
0.4	34.93	34.57	34.21	33.85	33.49	33.13	32.78	32.42	32.07	31.72
0.5	31.37	31.02	30.67	30.32	29.98	29.64	29.29	28.95	28.61	28.28
0.6	27.94	27.60	27.27	26.94	26.61	26.28	25.96	25.63	25.31	24.99
0.7	24.67	24.35	24.03	23.72	23.41	23.10	22.79	22.48	22.18	21.87
0.8	21.57	21.27	20.98	20.68	20.39	20.10	19.81	19.52	19.23	18.95
0.9	18.67	18.39	18.11	17.84	17.56	17.29	17.03	16.76	16.49	16.23
1.0	15.97	15.72	15.46	15.21	14.96	14.71	14.46	14.22	13.97	13.73
1.1	13.50	13.26	13.03	12.80	12.57	12.34	12.12	11.90	11.68	11.46
1.2	11.24	11.03	10.82	10.61	10.41	10.21	10.00	9.81	9.61	9.42
1.3	9.22	9.03	8.85	8.66	8.48	8.30	8.12	7.95	7.77	7.60
1.4	7.44	7.27	7.10	6.94	6.78	6.63	6.47	6.32	6.17	6.02
1.5	5.87	5.73	5.59	5.45	5.31	5.18	5.05	4.92	4.79	4.66
1.6	4.54	4.41	4.30	4.18	4.06	3.95	3.84	3.73	3.62	3.52
1.7	3.41	3.31	3.21	3.11	3.02	2.93	2.83	2.74	2.66	2.57
1.8	2.49	2.40	2.32	2.25	2.17	2.09	2.02	1.95	1.88	1.81
1.9	1.75	1.68	1.62	1.56	1.50	1.44	1.38	1.33	1.27	1.22
2.0	1.17	1.12	1.07	1.03	0.98	0.94	0.90	0.86	0.82	0.78
2.1	0.74	0.71	0.67	0.64	0.61	0.58	0.55	0.52	0.49	0.46
2.2	0.44	0.41	0.39	0.37	0.34	0.32	0.30	0.29	0.27	0.25
2.3	0.23	0.22	0.20	0.19	0.18	0.16	0.15	0.14	0.13	0.12
2.4	0.11	0.10	0.09	0.08	0.08	0.07	0.06	0.06	0.05	0.05
2.5	0.04	0.04	0.03	0.03	0.03	0.02	0.02	0.02	0.01	0.01
2.6	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00

NUMBERS IN BODY OF TABLE ARE ESTIMATES OF PERCENT DEFECTIVE CORRESPONDING TO SPECIFIC VALUES OF Q, THE QUALITY INDEX. FOR Q VALUES GREATER THAN OR EQUAL TO ZERO, THE PERCENT DEFECTIVE ESTIMATE MAY BE READ DIRECTLY FROM THE TABLE. FOR Q VALUES LESS THAN ZERO, THE TABLE VALUE MUST BE SUBTRACTED FROM 100.

FIGURE 1 A compact table for estimation of percent defective.

This situation can hardly inspire the confidence and cooperation of the construction industry, and it should cause the transportation profession more than a little chagrin. The necessary engineering knowledge exists to develop rational pay schedules that are fair and defensible. The method is based on the legal principle of liquidated damages, and the concept is simple and straightforward. The pay schedule is designed to withhold sufficient payment at the time of construction to cover the cost of future repairs made necessary by defective work. This is discussed in more detail in Part 3, and the complete development can be found in other recent publications (*15,16*).

Example 6: Importance of Operating Characteristic Curves

Perhaps the most common oversight in the development of SQA specifications is the failure to construct the operating characteristic curves to check whether in fact the acceptance plans will perform as desired. The following example is taken from a federal standard (*17*), and we understand that it will be included in the forthcoming WASHTO Model Quality Assurance Specifications. This particular generic acceptance procedure has some desirable features—it uses the more efficient standard deviation method and includes a bonus clause to reward superior performance—but it operates in a way that may be quite different from what most transportation agencies would consider appropriate.

In the text of the original standard, it is stated that the acceptable quality level (AQL) is defined as 95 percent within limits. The pay schedule provides for bonus pay factors up to 105 percent for still better quality and proportionally lower pay factors for lesser quality. However, to determine how the acceptance procedure will actually perform, it is necessary to construct the operating characteristic curves for the sample sizes that would be used. The operating characteristic curve in tabular form for this plan using a typical sample size of $N = 5$ is as follows:

Percent Within Tolerance	*Avg Pay Factor (%)*
100	105.0
95 (AQL)	103.7
90	102.2
85	100.5
80	97.8
75	94.3
70	89.9

It can be seen that the contractor who performs consistently at a quality level of 95 percent within limits, the level that has been defined as acceptable, will receive an average pay factor of about 104 percent. Furthermore, the contractor who furnishes work that is approximately 85 percent within limits, substantially below the level that has been defined as acceptable, will still receive an average pay factor of 100 percent. At a time when most states are scrambling just to make ends meet, a known and predictable overpayment such as this would be regarded by many as highly inappropriate.

For this particular acceptance plan to be considered satisfactory, the AQL would have to be redefined as 85 percent within limits and the transportation agency would have to be convinced that the rate of pay adjustment for other levels of quality was appropriate for the construction items to which it was applied.

POOR TEACHING OF STATISTICAL METHODS

As noted earlier in this paper, a cause for increasing alarm in this country has been a persistent decline in academic achievement in virtually all subjects, especially science and mathematics. In a presentation at the 1990 TRB Annual Meeting, C. V. Wootan of the Texas Transportation Institute summed it up as follows: "We are rapidly progressing from a nation at risk in our educational system to a nation in crisis. . . . The once unchallenged preeminence of America in science and technology is being overtaken by well educated, highly motivated, and determined competition throughout the world."

In his critical and highly acclaimed book *Out of the Crisis* (*5*), Deming claims that one of the root causes for the dismal performance of quality assurance practices in this country is the poor teaching of statistical methods. To paraphrase slightly, he notes that no one should teach SQA without a thorough knowledge of statistical theory through at least the master's degree level, supplemented by actual, hands-on experience under a master. He adds that this observation is made on the basis of experience, having witnessed countless examples of incompetent teaching and faulty application.

Having participated in SQA training as both students and teachers over a period of several years, we believe that Deming's observations are particularly relevant. Although the training in SQA that has been offered in the transportation field has been excellent in many respects (*18–20*), its one consistent shortcoming has been its failure to delve more than superficially into the theory underlying statistical acceptance sampling.

Why is this so important? The primary reason is the need to make quality assurance programs perform as effectively as possible. Practitioners with a minimal amount of training are often unaware that some methods are considerably better than others. Inferior methods simply do not perform well in terms of making correct decisions—accepting good quality and rejecting poor quality. The knowledge necessary to understand which methods are the best is not especially difficult to acquire, but unfortunately it usually is not included in elementary SQA courses.

Although it might seem at first that almost anyone with a smattering of statistical knowledge could teach a beginners' course in SQA, a little reflection suggests why Deming considers such an arrangement unacceptable. SQA, like any mathematical discipline, requires logical and rigorous thinking. Because many beginners taking their first SQA course have not yet acquired this ability, it it vital that the first exposure be with an experienced instructor capable of instilling the necessary critical thinking skills. Once a sense of technical rigor is ingrained, it becomes almost instinctive to understand the subject and apply it correctly. In our opinion, the only reason SQA has the reputation of being a difficult subject is because it has not been taught as well as it might be.

Another important reason why top-level instruction is essential is the belief that open and inquiring minds are among the nation's most valuable resources. Some of the people attending SQA courses will be future managers and leaders who may someday have to make policy decisions on these matters. It is in the best interest of the transportation profession to provide them with a thorough understanding of these principles so that their future decisions will be wisely guided.

And finally, credibility is an important issue. The construction industry is often suspicious and distrustful of SQA methods, and this tends to produce an almost adversarial relationship when these procedures are introduced. To overcome this, transportation agencies must be prepared to quickly and forthrightly provide correct answers and plausible explanations for a wide variety of technical questions. Any delay or inability to answer will be interpreted as evasive. In this case, thorough training as advocated by Deming could be instrumental in fostering a more cooperative and productive relationship.

RESISTANCE FROM WITHIN THE TRANSPORTATION AGENCY

The adoption of a full-fledged quality assurance program will often be a significant departure from procedures with which transportation agency personnel are familiar and comfortable. For those who are accustomed to exercising control over the construction process through method-type specifications, it may be difficult to get used to the broad latitude of control given to contractors under end-result specifications. This requires new thinking, new inspection and acceptance techniques, and a willingness to adapt to the new procedures.

Some degree of resistance is only natural. It is not reasonable to expect widespread support for any new program until the reasons for its adoption are understood. Management that clearly sees the advantages of a quality assurance program is in a much better position to communicate this to the rest of the organization through a well-conceived action plan. This is discussed in more detail in Part 4.

The first step is to conduct a series of training sessions to acquaint managers and specification writers with the underlying logic and rationale. The training need not be highly theoretical and should cover such basic topics as normal process variation, the estimation of construction quality by random sampling, and the development of effective statistical acceptance procedures. The training should be gradually expanded to include construction and materials personnel, stressing the important role they play in implementing quality assurance specifications. And finally, it may be helpful at some point to open the training sessions to a limited number of contractors' personnel so that they too will develop an understanding of the basic requirements and how to meet them.

Ideally, a few engineers should receive some additional, college-level instruction so that at least one individual in the organization will have a thorough understanding of the statistical concepts involved. To provide the necessary combination of skills to support the NJDOT quality assurance program, an entirely new set of job specifications was created—the Statistical Engineer Series. A description of these positions is provided in Part 4.

At NJDOT another activity was found to be extremely helpful in overcoming internal resistance. Field staff, who deal with the new procedures on a daily basis, often have very useful suggestions to make concerning what works, what

doesn't, and what might be improved. Although there are procedures for formally transmitting such suggestions upward through the chain of command, there are also various reasons why such suggestions often become distorted or lost in the process. In the course of developing the specification described in the case study in Part 1, a group of field staff was invited to come in and provide direct feedback to both top and middle management on any implementation problems they were having. The meeting was very successful in that it promoted a healthy dialogue and generated several suggestions, some of which were adopted. The participants left with the knowledge that their concerns had been heard, that they had played an important role in developing the specification, and that there were valid reasons why certain changes could not be made. This type of participation not only improved the specification but also improved the morale of the field staff.

Yet another general approach appears to be extremely effective whenever a new specification is implemented. The new procedure is first tested by simulation on field data from several ongoing jobs. Next, a partial implementation on one or two small projects is scheduled with certain provisions tempered slightly, such as reducing all pay adjustments by half. Finally, provided the pilot projects have been successful, a full-scale implementation is undertaken. Staging the implementation in this manner seems to be more acceptable to all concerned. It allows for a period of learning and adjustment for both state and contractors' personnel, and if there should be any technical or administrative difficulties, they can be corrected with a minimum amount of inconvenience.

OPPOSITION FROM THE CONSTRUCTION INDUSTRY

Virtually every agency that has adopted SQA methods has had to overcome considerable opposition from the construction industry. Although much of the resistance can be attributed to a general fear of the unknown, a substantial part has been due to the manner in which statistical specifications have evolved. Over the years, contractors have been forced to suffer through the growing pains of this discipline as statistical specification writers were learning their craft by trial and error. The state of the art has progressed considerably in recent years, and it is now possible to write quality assurance specifications that are simple, clear, technically sound, and fair to the seller while protective of the interests of the buyer. In our opinion, this has resulted in a gradual softening of the resistance that was typically encountered. If this progress can be continued to the point at which a more thorough, consistent, and scientific approach is adopted nationwide, it is conceivable that the construction industry might eventually come around to endorsing these methods, at least in a qualified way. A recent article (*21*) by an officer of a large highway construction firm indicates that contractors have begun to discover that SQA methods offer benefits to them as well.

POLITICAL FACTORS

Engineering decisions are often based on factors other than the purely technical. When the various engineering alternatives are approximately equivalent, this is both practical and desirable. But when there is a considerable difference in terms of performance and cost-effectiveness, decisions based on political considerations can sometimes work to the disadvantage of society as a whole. Such factors may be responsible for the failure of the transportation profession to take action that would be in its best interest, particularly in the area of quality assurance. For example, some administrations and standards-writing groups may be reluctant to phase out ineffective practices because of a fear that this would be interpreted as an admission of prior misfeasance.

Furthermore, administrators may view the establishment of sound quality assurance practices as having little political payback because there are few obvious short-term benefits. In fact, the immediate effects are often unfavorable, such as the resistance that is typically encountered from the construction industry. The real benefits, in terms of reduced maintenance and extended service lives, will not be realized until many years later.

Unfortunately, the converse of this is also true. An administration might be inclined to skimp on quality assurance during times of tight budgets, realizing that the real effects of this false economy are not likely to be felt for years. A preferred response to tight budgets would be to "work smarter" by using the most effective procedures available.

It is incumbent upon responsible leaders to put aside concerns of possible short-term criticism or failure to achieve immediate political gain in favor of the much greater long-term benefits to be derived from the adoption of technically sound, state-of-the-art practices. This decision could be made easier for individual agencies if it were carried out as part of a joint national effort. The establishment of a national policy on quality assurance, as advocated in Part 1 of this paper, would provide an ideal way to accomplish this.

WORK ETHIC AND CULTURAL ATTITUDES

The following example from Deming (*5*) is both incisive and revealing. William G. Ouchi, after observing the participants at a meeting of an American trade association adjourn early each day for a variety of recreational activities, commenced his presentation as follows:

> While you are out on the golf course this afternoon, waiting for your partner to tee up, I want you to think about something. Last month I was in Tokyo, where I visited your trade association counterpart. It represents the roughly two hundred Japanese companies who are your direct competitors. They are now holding meetings from eight each morning until nine each night, five days a week, for three months straight, so that one company's oscilloscope will connect to another company's analyzer, so that they can agree on product safety standards to recommend to the government (to speed up getting to the marketplace), so that they can agree on their needs for changes in regulation, export policy, and financing and then approach their government with one voice to ask for cooperation. Tell me who you think is going to be in better shape five years from now.

One must seriously question whether the work ethic and level of commitment in the transportation field are capable of meeting today's challenges with respect to quality assur-

ance. Do transportation "trade association" meetings (those of AASHTO, TRB, etc.) suffer from any of the shortcomings noted by Ouchi? Are transportation professionals properly assessing the level of effort required to turn out standards and specifications of a high caliber, or are they willing to accept whatever can be hammered out in an hour or two at a committee meeting? Is the necessary expertise being brought to bear on these issues, or are transportation committees and agencies willing to settle for whatever talent happens to be readily available? Have transportation engineers insisted that all papers, standards, and specifications reflect a high level of technical competence, or have they been willing to accept whatever they get? Have the necessary formal review processes been established to ensure that technical standards receive a critical, independent evaluation, or has there been a willingness to gamble that the balloting process will uncover any deficiencies?

These are telling questions and, in our opinion, candid answers will reveal a process that is almost guaranteed to produce mediocre results. The necessary technology exists, the talent to use that technology exists, but the process that would ensure the proper use of that talent and technology does not exist.

It might seem that radical cultural changes would be necessary to dispell the old notion of "close enough for highway work" and replace it with an ingrained desire for excellence. Juran, however, argues for a different approach (*6*). He notes that, in actual practice, it often is first necessary to bring about behavioral change in the form of new procedures and, after the new procedures have been demonstrated to be effective, the desired change in attitude will follow.

We believe that just such an approach is required to resolve the problems with transportation quality assurance. Leaders must first address the questions raised earlier in this section and then create a process that both fosters and demands excellence. This will produce improved standards and specifications and these, in turn, will ultimately provide the strongest testimonial to their own worth.

CONVEYING THE WRONG MESSAGES

Two of us recently attended a technical advisory meeting that had been convened to address the status of transportation quality assurance. The meeting began with a series of presentations to focus the group's attention on several specific issues. The participants were then divided into smaller subgroups to brainstorm how these issues might be addressed. The meeting concluded with a series of summary presentations.

In one of the summary presentations, the speaker noted with an apparent sense of accomplishment that he had covered the topic of quality assurance without once mentioning statistics. Although it probably was not intended that way, such a remark could imply that a thorough understanding of statistical principles is not a prerequisite for an effective quality assurance program. In our opinion, this is not the message that should be conveyed, nor is it the message found in the TQM literature and the recent documentaries on TQM on the public broadcasting channels.

What, then, is the message that *should* be conveyed? We believe that the following points should be stressed to both state and contractors' personnel alike:

1. It will be necessary to learn some new things, particularly in the areas of elementary statistics, process variability, and acceptance sampling.
2. Some formal training will be required. Something more than a single college-level course will be required for at least one member of an organization.
3. And finally, if a conscientious effort is made to understand and apply quality assurance technology, this will more than pay for itself in terms of quality achieved, reduction of rejections and rework, and a generally smoother-running operation.

There are still other ways in which the wrong message is sometimes conveyed. Management that does not visibly and actively support excellence in all aspects of engineering is communicating in a subtle way to its employees that excellence is not really that important. When organizations such as AASHTO publish SQA standards that are technically unsound, the message communicated to the transportation community at large is that it really is not essential that scientific principles be applied correctly. When the transportation profession as a whole is willing to tolerate the great disparity with which quality assurance is applied across the country, the message communicated to the construction industry is that the profession is either unconcerned about establishing consistent and valid practices or else is incapable of doing so.

If there is to be any chance of reversing what appears to be a long-term slide into technical mediocrity, it will be up to transportation leaders to begin communicating some distinctly different messages, much like those found in the writings of Crosby (*22*), Deming (*5*), and Juran (*6*). Although the advice of these authors has been directed primarily at the private sector, a substantial portion of it is equally applicable in the public sector. It is now up to transportation leaders to read it, assimilate it, and begin applying it.

SUMMARY AND PREVIEW

Several obstacles—technical, managerial, political, and cultural—could impede either the effective application of SQA procedures by individual agencies or the establishment of a sound national policy on quality assurance. Perhaps the biggest obstacles are a general lack of awareness of just how far behind the state of the art transportation quality assurance really is, the failure to insist that those involved with quality assurance be thoroughly educated in these matters, and a work ethic that seems to be devoid of any real pride in what it produces. Many of these issues have recently been addressed by TQM writers such as Crosby, Deming, and Juran, and transportation leaders need to become more familiar with their work and seek to apply it in the public sector. As recommended in Part 1, a national task force of "can do" leaders should be created to focus a multidisciplinary attack on these problems.

If the full potential of quality assurance techniques is to be realized, a number of basic concepts must be understood and

applied correctly. To aid in this effort, Part 3 presents a series of fundamental principles that underlie the type of acceptance sampling most suited for transportation applications. Part 4 then outlines a plan of action to effectively apply the concepts developed in the previous three sections of the paper.

REFERENCES (Part 2)

1. R. M. Weed. Revision of a Flawed Acceptance Standard. In *Transportation Research Record 1056*, TRB, National Research Council, Washington, D.C., 1986, pp. 21–35.
2. R. M. Weed. Adjusted Pay Schedules: New Concepts and Provisions. In *Transportation Research Record 986*, TRB, National Research Council, Washington, D.C., 1984, pp. 32–38.
3. R. M. Moore, J. P. Mahoney, R. G. Hicks, and J. E. Wilson. Overview of Pay Adjustment Factors for Asphalt Concrete Mixtures. In *Transportation Research Record 821*, TRB, National Research Council, Washington, D.C., 1981, pp. 49–56.
4. P. A. Kopac. Current Practices in Acceptance of Bituminous Concrete Compaction. In *Transportation Research Record 986*, TRB, National Research Council, Washington, D.C., 1984, pp. 43–46.
5. W. E. Deming. *Out of the Crisis*. Massachusetts Institute of Technology, Cambridge, 1986.
6. J. M. Juran. *Juran on Leadership for Quality*. The Free Press, New York, 1989.
7. E. D. Lieberman. *Unfit to Manage*. McGraw-Hill Book Company, New York, 1988.
8. M. Landsberg. Interstate Almost Finished after 35 Years. *Trenton Times*, June 16, 1991, Section BB, p. 1.
9. Bridge Inventory Update. *Better Roads*, Nov. 1990, p. 26.
10. M. Maccoby. Deming Critiques American Management. *Research Technology Management* (Industrial Research Institute, Inc., Washington, D.C.), May-June 1990, p. 43.
11. R. Thomas and E. Thomas. Where Did All the Money Go? *Newsweek*, July 1, 1991, pp. 25–28.
12. *Recommended Practice for Evaluation of Strength Test Results of Concrete*. ACI Standard 214-77. American Concrete Institute, Detroit, Mich., 1977.
13. Standard R9-86: Acceptance Sampling Plans for Highway Construction, Method B. In *AASHTO Standard Specifications for Transportation Materials and Methods of Sampling and Testing*, Part I. AASHTO, Washington, D.C., 1986.
14. *Advancement of SQA and TQM in Design and Construction*. Study 7490. New Jersey Department of Transportation, Division of Research, Trenton, in preparation.
15. R. M. Weed. *Statistical Specification Development*, 2nd ed. Final Report, New Jersey Department of Transportation Study 7788. NJDOT, Trenton, 1989.
16. R. M. Weed. Method to Establish Pay Schedules for Rigid Pavement. In *Transportation Research Record 885*, TRB, National Research Council, Washington, D.C., 1982, pp. 18–24.
17. *Standard Specifications for Construction of Roads and Bridges on Federal Highway Projects*. FP-85, Section 106. FHWA, U.S. Department of Transportation, 1985.
18. J. H. Willenbrock. *Statistical Quality Control of Highway Construction*, Vols. I and II. FHWA, U.S. Department of Transportation, 1976.
19. *Demonstration Project 42: Highway Quality Assurance*. Demonstration Projects Division, FHWA, U.S. Department of Transportation, 1976.
20. J. L. Burati and C. S. Hughes. *Highway Materials Engineering, Module I, Materials Control and Acceptance—Quality Assurance*. National Highway Institute, FHWA, U.S. Department of Transportation, 1990.
21. M. L. Powell. Quality—Contractors Cannot Afford Less. *TR News*, Nov.-Dec. 1986, pp. 38–43.
22. P. B. Crosby. *Quality Is Free*. McGraw-Hill Book Company, New York, 1979.

PART 3: FUNDAMENTAL CONCEPTS

The purpose of Part 3 is to offer a series of fundamental principles as a starting point for the development of a consistent and sound national policy on transportation quality assurance. The concepts that follow have evolved and been proven by actual field application over a period of approximately 20 years.

The following sections deal with the fundamental concepts that pertain to the development of technically sound, defensible specifications:

1. Objectives of SQA specifications,
2. Relationship between quality and performance,
3. Choice of appropriate statistical parameter,
4. Acceptable and rejectable quality levels,
5. Attributes and variables plans,
6. Operating characteristic curves and risk analysis,
7. Computer simulation,
8. Lot sizes and sample sizes,
9. Random sampling procedures,
10. Basis for pay adjustments,
11. Justification for bonus clauses,
12. Advantage of continuous pay schedules, and
13. Development of acceptance plan.

Some sources of possible confusion that could prevent transportation agencies from realizing the full potential that SQA has to offer are discussed in the following sections:

14. Accuracy, precision, and bias;
15. Sample coverage and quality assurance;
16. Averaging power of statistics;
17. Appropriate measures of variability;
18. Minor problem with variables plans;
19. Problem with variability known procedures;
20. Dual acceptance procedures;
21. Retesting provisions;
22. Zero pay factors;
23. Availability of statistical tables;
24. Classical versus Bayesian methods;
25. Moving averages;
26. Unbalanced bids;
27. Non-pay-adjustment items; and
28. Legal considerations.

1. OBJECTIVES OF SQA SPECIFICATIONS

For any program to be successful, there must be a clear understanding of the objectives. We consider the following objectives to be most important:

1. The primary objective is to communicate to the contractor in a clear and unambiguous manner exactly what is wanted. Various statistical measures provide a practical and convenient way to describe the desired end result.
2. In keeping with the end-result philosophy, the contractor should be given most of the responsibility for controlling the construction process, whereas the specifying agency should be primarily responsible for judging the acceptability of the finished work.
3. There should be sufficient incentive for the contractor to produce the desired quality (or better). This can be accomplished by means of adjusted pay schedules, which assess pay reductions for deficient quality and, when appropriate, award suitable bonuses for superior quality.
4. Ideally, the specification should pay 100 percent, on average, for acceptable work, and it should be fair and equitable in assigning pay factors for work that differs from the desired quality level.
5. The specification should be realistic in defining acceptable quality levels (AQLs) and rejectable quality levels (RQLs). The AQLs should be set high enough to satisfy design requirements but not so high that extraordinary methods or materials will be required. The RQLs should be set low enough that, when they occur, the option to require removal and replacement is truly justified.
6. It should be clear to the contractor what the appropriate target level of quality must be in order to receive 100 percent payment.

2. RELATIONSHIP BETWEEN QUALITY AND PERFORMANCE

In general, any construction specification should be applied to those parameters that are believed to be strongly related to the ultimate performance of the final product. In most cases, the qualitative relationship between commonly measured construction characteristics and performance has been well established. For example, compressive strength is known to be highly correlated with the performance of concrete structures even though the exact nature of this relationship may be somewhat vague. Furthermore, the relationship is a consistent one because at any reasonable level of strength, an increase in strength will provide still better performance.

In most cases, more than one quality characteristic must be considered. In the example just cited, concrete compressive strength alone may be insufficient to ensure the desired performance. To be durable, a concrete bridge deck must also have the necessary amount of entrained air. For concrete pavement, adequate thickness as well as strength must be achieved. For bituminous pavement, several requirements must be met simultaneously. It is the responsibility of the developer of the specification to include all important variables in a way that is logical and appropriate.

3. CHOICE OF APPROPRIATE STATISTICAL PARAMETER

Although various statistical measures of quality are available, transportation engineers have exhibited a strong preference for the concept of lot percent defective, the estimated percentage of the lot falling outside specification limits (or its counterpart, the percent within limits). This measure is particularly appealing for at least three reasons:

1. It can be applied to virtually any construction quality characteristic.
2. It encourages uniformity in that it controls both the average level and the variability of the product in a statistically efficient way.
3. Uniform quality, consistently within specification limits, is believed to be strongly associated with ultimate performance.

Figure 1 shows the concept of percent defective applied to both single-limit and double-limit specifications. The engineering rationale underlying this concept is discussed in American Concrete Institute (ACI) Standard 214 (*1*), for example, and is believed to be valid for a broad range of engineering applications.

4. ACCEPTABLE AND REJECTABLE QUALITY LEVELS

The AQL is the level of quality, usually defined in terms of some minimal degree of deficiency, that the specifying agency

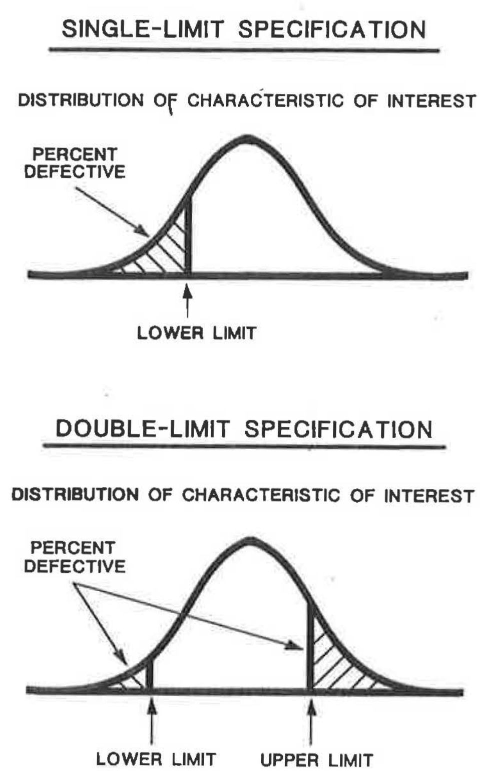

FIGURE 1 Concept of percent defective.

is willing to accept at 100 percent payment. The RQL is the level of quality that is so deficient that either repair or removal and replacement may be necessary. In between the AQL and the RQL, the work is considered to be marginally satisfactory and is accepted at reduced payment. (If, for practical reasons, an RQL item is left in place, it typically is assigned some minimum pay factor.)

Appropriate definitions of AQL and RQL are of considerable importance because this concept has far-reaching consequences. If the AQL is set at an unrealistically high level, the cost of such exceptional quality may far exceed its value to the transportation agency. At the other extreme, too low an AQL may produce lower bid prices, but the acceptance of consistently lower levels of quality may result in greatly increased maintenance costs and be more expensive in the long run. It is incumbent upon the transportation agency to exercise careful judgment in selecting AQL values that properly balance quality and economy in a realistic manner.

Whereas it is the prerogative of the transportation agency to define the AQL at any level it considers appropriate, there is somewhat less latitude in defining the RQL. Because of the severe consequences imposed upon the contractor when an item of RQL quality is detected, there may be litigation if the RQL is set at a level that does not clearly warrant such drastic action. Conceptually, at least, this poses no great problem. Since it is the purpose of the adjusted pay schedule to recoup losses expected from poor-quality work, it can be relied upon to perform its function down to some low level of quality that is defensible as the definition of the RQL.

Although the relationship between quality and performance of a construction item is usually known only in a vague, imprecise way, theoretical considerations can still provide some insight to aid in the selection of realistic AQL and RQL values. For example, ACI Standard 214 takes the following probabilistic approach regarding concrete compressive strength:

> If a small percentage of the test results fall below the design strength, a corresponding large percentage of the test results will be greater than the design strength with an equally large probability of being located in a critical area (of the structure). (*1*)

This clearly implies that some small percentage of strength tests falling below the design strength can be tolerated. ASTM C-94 goes even further to suggest the following:

> For concrete in structures designed by the ultimate strength method and in prestressed structures, not more than 10 percent of the strength tests shall have values less than the specified strength. (*2*, p. 65)

ASTM C-94 also goes on to state that as much as 20 percent of the strength tests may be below the specified strength for concrete designed by the working stress method. Standards such as these have been quite helpful and have led highway agencies to typically establish AQL values in the range of 5 to 10 percent defective.

ACI Standard 214 and ASTM C-94 have not addressed the question of defining the RQL, however, since the acceptance procedures that they advocate are strictly pass-or-fail methods. Anything that is not AQL is considered to be RQL. It is the concept of adjusted payment, which recognizes that there is an intermediate zone between clearly good and clearly poor quality, that requires distinct definitions of AQL and RQL.

When available, historical data can provide the basis upon which suitable definitions of AQL and RQL can be based. To be valid and useful, the data should have been gathered in a random fashion from a variety of projects spanning as broad a range of quality as possible. For example, if a highway agency found that pavements typically had performed satisfactorily when no more than 10 percent of the tests were less than the design thickness and it was desired to develop a specification that would continue to produce this same level of quality, then the AQL could be defined as 10 percent defective. Similarly, if very troublesome maintenance problems were found to be associated with a percent defective level of approximately 50 percent, then this might be an appropriate RQL.

Although these particular values for AQL and RQL might be reasonable, they are presented here only as examples to suggest the type of association between quality and performance that should be sought from historical data. When such an analysis is actually performed, it is necessary to screen out the effects of other contributing factors such as the strength of the pavement layer. Finally, when the AQL and RQL are derived in an empirical manner such as this, some form of long-term monitoring of the acceptance procedure may be desirable. The random sampling plans will continue to produce valid historical data that can be used to review the effectiveness of the specification at some future date.

5. ATTRIBUTES AND VARIABLES PLANS

Percent defective (or its counterpart, percent within limits) can be controlled by either of two types of acceptance procedure—attributes plans or variables plans. Attributes plans typically involve the counting of some type of defect, or the number of failing tests, and lead to the classification of the inspected lot as either satisfactory or unsatisfactory. Variables plans apply to quality characteristics that are measured on a continuous scale and involve the computation of statistical parameters such as the mean and standard deviation. Either type of plan may be used for pass-or-fail decisions, but variables plans are somewhat more convenient as a basis for adjusted pay schedules.

As a general rule, variables plans are more discriminating than attributes plans. This means that, for a given sample size, greater protection is provided or, for a given level of protection, a smaller sampling effort is required. Either way, substantial economic benefits can be realized with the use of variables plans. This is demonstrated by the examples in Section 13.

A basic assumption of variables acceptance theory is that the population (lot) being sampled is normally distributed. Many construction characteristics have been found to closely approximate the normal distribution, thereby justifying the widespread use of variables procedures. When a situation occurs in which the construction characteristic is distinctly nonnormal, there are two possible remedies. Either the individual tests can be replaced with the averages of two or more tests, a step that greatly improves normality, or an at-

tributes procedure can be used, which requires no distributional assumptions.

Variables plans can use either the standard deviation or the range as the measure of variability. In the past, the range was often used because it was easier to understand and compute. Today, the standard deviation is a more commonly used statistical measure because it makes more efficient use of the available data. Just as variables plans are more discriminating than attributes plans, standard deviation plans are generally superior to range plans. An illustration of this is given in Section 17.

The book by Duncan (*3*) is an excellent general reference on attributes and variables sampling, and some examples are given in Section 13 of this paper. A recent publication (*4*) provides certain tables in a more convenient form.

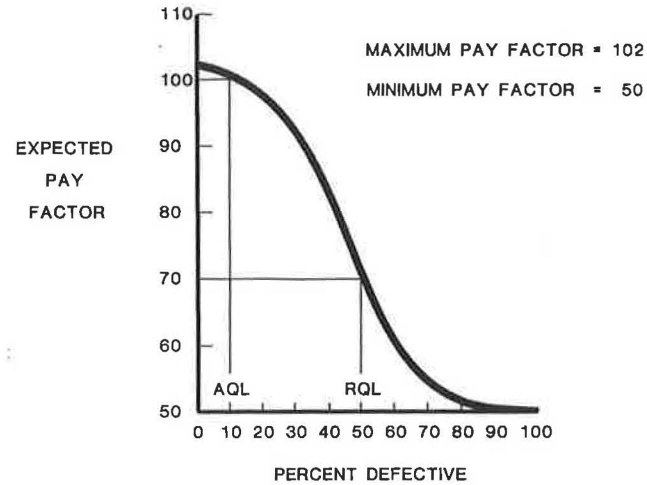

FIGURE 3 Typical OC curve for statistical acceptance procedure with an adjusted pay schedule.

6. OPERATING CHARACTERISTIC CURVES AND RISK ANALYSIS

An absolutely vital step in the development of a statistical specification is the construction of the operating characteristic (OC) curve. This is the only way to know in advance whether or not the acceptance procedure will function as intended. It is through the study of such curves that the risks to both parties can be recognized and controlled at suitably low levels. This enables the highway agency to develop fair and effective specifications and may aid the contractor in determining the appropriate bidding and production strategies.

A conventional OC curve is shown in Figure 2. Probability of acceptance is indicated on the *Y*-axis for the range of quality levels (indicated schematically in this example) on the *X*-axis. The contractor's risk of having good (AQL) material rejected and the agency's risk of accepting poor (RQL) material are both illustrated.

Figure 3 presents an OC curve constructed for a statistical specification with an adjusted pay schedule. Quality levels are indicated on the *X*-axis in the usual manner, but, instead of probability of acceptance, the *Y*-axis gives the expected pay factor.

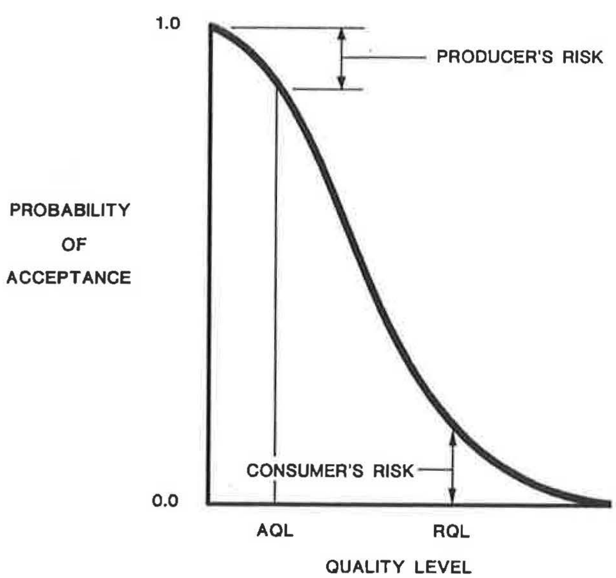

FIGURE 2 Conventional OC curve.

Although the risks have a slightly different interpretation when associated with the expected payment curve in Figure 3, essentially the same type of information is provided. In this particular example, AQL work receives an expected pay factor of 100 percent, as desired. At the other extreme, RQL work corresponds to an expected pay factor of 70 percent. Presumably the transportation agency has determined that this will cause sufficient money to be withheld to cover the anticipated cost of future repairs. For still lower levels of quality, the curve levels off at the minimum pay factor of 50 percent.

In the case of pass-or-fail acceptance procedures, OC curves of the type shown in Figure 2 can be computed directly or constructed with the aid of special tables such as those presented in Section 23. For acceptance procedures with adjusted pay schedules, OC curves of the type shown in Figure 3 can be obtained by using special software (*5*) or by developing relatively simple computer simulation programs.

7. COMPUTER SIMULATION

Computer simulation is one of the most powerful analysis methods available for handling a wide variety of complex problems and yet, contrary to what might be expected, it is one of the simplest to understand and apply (*6,7*). The ease of this approach appeals both to problem solvers and to those to whom the results of an analysis must be presented and explained. Most simulations require only the following basic steps:

- Generate random data simulating the real process,
- Apply the procedure that is to be tested, and
- Observe the results.

We routinely use computer simulation to test new statistical acceptance procedures before their actual implementation. A typical run is shown in Figure 4, in which it was desired to test a pay schedule for concrete compressive strength on 1,000

```
RUN SPECSIM
EXECUTION BEGINS...


ENTER COEFFICIENTS A AND B OF PAY EQUATION PF = A - B(PD)
?
102 0.2

ENTER SPECIFICATION LIMIT AND SAMPLE SIZE
?
4000 5

ENTER STANDARD DEVIATION (PRODUCT AND TESTING VARIABILITY)
?
300

ENTER PERCENT DEFECTIVE LEVEL AND RANDOM GENERATOR SEED
NUMBER
?
10 123456789


DISTRIBUTION OF SIMULATED TEST RESULTS
     N = 5000
     MEAN = 4382
     STANDARD DEVIATION = 305
     PERCENT DEFECTIVE = 10.5


DISTRIBUTION OF PD ESTIMATES
     N = 1000
     MEAN = 10.6


AVERAGE PAY FACTOR = 99.9
```

FIGURE 4 Computer simulation test of a statistical acceptance procedure.

simulated lots at a quality level of 10 percent defective. In this example, it is seen that the computer generated a total of 5,000 random test results at a percent defective level of $PD = 10.5$, very close to the desired value of 10.0. The average estimated percent defective for the 1,000 lots was 10.6, also very close to the true value of 10.5, as would be expected of a valid statistical estimation procedure. For this example, the AQL is considered to be 10 percent defective and, accordingly, the average pay factor is almost exactly 100 percent, indicating that the acceptance procedure is performing properly at this quality level.

A simulation program such as this, because it uses several previously developed subroutines, requires only about an hour to prepare and enter on the computer. Whereas earlier efforts of this type were done primarily on mainframe computers, the increasing availability of Fortran compilers has now made it possible to do much of this work on relatively inexpensive personal computers. A run such as that shown in Figure 4 might require a fraction of a second on a mainframe and only a few seconds on a PC. The Fortran coding for a variety of useful simulation subroutines may be found in a recent report (*8*).

By using computer simulation in this manner, it is possible to ensure that statistical specifications will protect the interests of the transportation agency and also that they will be fair to the construction industry. This approach also provides an effective way to acquaint contractors with the degree of control that must be achieved to meet the requirements of a new specification.

8. LOT SIZES AND SAMPLE SIZES

The selection of lot size is dictated primarily by practicality and convenience, although, for variables acceptance procedures, care must be exercised in combining work produced at different times or under different conditions because this might violate the assumption of normality. Typically, either time or quantity limits are used to define lots, such as a day's production or 5,000 yd^2.

The sample size is a more important consideration because this has a direct effect on the risks involved. Except for attributes sampling from discrete lots (items that are counted), the lot size plays no role in the development of the OC curve. Usually, but not always, larger sample sizes reduce the risks to both the contractor and the transportation agency, but to be sure the plan will perform as desired, the OC curves should be constructed for all sample sizes under consideration.

Attributes plans may be used with sample sizes as small as $N = 1$, although a plan with such a small sample will obviously be quite weak. For mathematical reasons, variables plans require a sample size of $N = 3$ or larger.

In general, for a given sample size and level of protection, defining larger lot sizes will reduce the overall level of effort devoted to sampling and testing. Assuming that larger lot sizes are satisfactory from a statistical standpoint, a price may still have to be paid for this economy. If there were a problem with the quality being produced, a proportionally larger quantity of defective material would have been produced before the problem was discovered. For this reason, lot sizes in the transportation field tend to be relatively small, seldom including more than a day's production.

Although we don't advocate such an approach, there is no theoretical reason why an entire project cannot be treated as a single lot. If it is known that long-term production closely approximates a normal distribution, a variables procedure could be used to make more efficient use of the data. If long-term production cannot confidently be assumed to be normal, then an attributes procedure would be necessary. A sufficiently large sample would be required to be confident that this single decision on the entire project was correct. Examination of the OC curve might suggest the desirability of a sequential acceptance procedure under which, if the first set of tests did not provide a clear-cut accept or reject decision, a second set of tests would be obtained to clarify the issue.

9. RANDOM SAMPLING PROCEDURES

Of the various theoretical assumptions upon which statistical acceptance procedures are based, random sampling is one of the most important. Only when all vestiges of personal bias have been removed can the laws of probability be relied upon to function properly.

Random sampling is often defined as a manner of sampling that allows every member of the population (lot) to have an equal opportunity of appearing in the sample. Stratified random sampling, which requires the drawing of a single sample from each of a number of equal-sized sublots, satisfies this basic requirement and guarantees that the samples will never be clustered in one small portion of the lot. When the product to be sampled can be measured on a continuous scale, this procedure is relatively straightforward. Figure 5 shows a stratified random sampling procedure applied to highway pavement.

```
            (NUMERALS 1 - 5 INDICATE SAMPLING LOCATIONS)

          1000'       1000'        1000'       1000'       1000'
 24' +-----------+-----------+-----------+-----------+-----------+
  :  :     1     :           :           :           :     5     :
 Y : :           :           :3          :           :           :
  :  :           :          2:           :  4        :           :
  0  +-----------+-----------+-----------+-----------+-----------+

     0 -------------------------------------------------------- 5000'
                                 X
```

DETERMINATION OF RANDOM X COORDINATES

SAMPLE NUMBER	RANDOM NUMBER	MULTIPLICATION TERM (SUBLOT LENGTH)	ADDITION TERM (CUMULATIVE LENGTH TO THIS SUBLOT)	X
1	0.603	x 1000	+ 0	= 603
2	0.992	x 1000	+ 1000	= 1992
3	0.086	x 1000	+ 2000	= 2086
4	0.214	x 1000	+ 3000	= 3214
5	0.551	x 1000	+ 4000	= 4551

DETERMINATION OF RANDOM Y COORDINATES

SAMPLE NUMBER	RANDOM NUMBER	MULTIPLICATION TERM (PAVEMENT WIDTH)	Y
1	0.750	x 24	= 18
2	0.286	x 24	= 7
3	0.542	x 24	= 13
4	0.081	x 24	= 2
5	0.877	x 24	= 21

FIGURE 5 Stratified random sampling procedure applied to highway pavement.

In other situations, however, it is more convenient to measure the product in discrete units such as items, batches, or truckloads. For these cases also, it is desirable to have a sampling procedure that performs like the continuous stratified procedure, that is, one that spreads the samples throughout the lot while allowing each item an equal opportunity of being included in the sample. Figure 6 presents a worksheet that we have found especially useful. Once the procedure has been demonstrated, the instructions at the bottom are all that is necessary to guide the user. A more complete description of the steps is as follows:

1. For the example shown in Figure 6, three trucks are to be randomly selected from a total lot size of 20 trucks delivering material to the job site.
2. In the sample selection table (upper table), all numbers greater than the lot size are crossed out (numbers 21–100).
3. Next, the remaining numbers (1–20) are divided into three approximately equal subgroups, one for each sample that is to be taken, by underlining them (1–6, 7–13, 14–20).
4. Following this, a random sample is chosen from each subgroup. The user is instructed to touch some location in the random number table (lower table) with the point of a pencil without looking and, moving from there in a predetermined manner, to continue until a number within the first subgroup (1–6) is obtained. This number is circled in both the random number table and the sample selection table. This procedure is then repeated for the second and third subgroups. In this particular example, samples are to be taken from the second, eleventh, and fifteenth trucks arriving at the job site.

A careful analysis of this procedure would show that each truck in the total lot of 20 trucks has very nearly, but not exactly, the same probability of appearing in the sample. This slight imperfection is the result of the unequal subgroup sizes. If it were believed necessary to overcome this minor departure from pure randomness, this could be accomplished by choosing a random starting point for the subgrouping rather than beginning with the first truck. A variety of procedures of this type have been developed and are described in other publications. (*8*,*9*).

10. BASIS FOR PAY ADJUSTMENTS

The major concern in the development of adjusted pay schedules is the determination of appropriate pay levels for various levels of quality. Over the years, several methods have been proposed (*10–15*), and when there was little or no information relating quality measures to performance, the methods were necessarily quite arbitrary. In cases for which quality-performance relationships have been established (or can be estimated), one of the more rational methods for developing pay schedules is based on the legal principle of liquidated damages. In this approach, the pay schedule is designed to withhold sufficient payment at the time of construction to cover the cost of future repairs made necessary by defective work.

The complete development of the liquidated-damages approach can be found in either of two recent publications (*8*, *13*) and will be described only briefly here. In the case of highway pavement, for example, the thickness and material characteristics are chosen to carry the estimated loading for the desired service life. At the end of its useful service life, the pavement will commence receiving a series of overlays that, based on our experience, typically last about 10 years. If because of construction deficiencies the pavement is not capable of carrying the design loading, it will fail prematurely. When this happens, the series of overlays that follow the initial design period will be moved forward in time, resulting in an extra expense to the transportation agency. Using engineering economics principles, it is possible to develop an expression giving the appropriate pay factor for various levels of expected life, such as that given by Equation 1.

$$PF = 100[1 + C_o(R^{L_d} - R^{L_e})/C_p(1 - R^{L_o})] \quad (1)$$

where

PF = appropriate pay factor (percent),
C_p = present unit cost of pavement (bid item only),
C_o = present unit cost of overlay (total in-place cost),
L_d = design life of pavement,
L_e = expected life of pavement,
L_o = expected life of overlay,
R = $(1 + R_{inf}/100)/(1 + R_{int}/100)$,
R_{inf} = annual inflation rate (percent), and
R_{int} = annual interest rate (percent).

To illustrate the degree of pay adjustment that this approach produces, the following typical values have been assumed:

C_p = \$40/yd^2 (typical bid for NJDOT-design concrete pavement),

ROUTE _______________ SECTION ______________________ DATE ____/____/____ LOT ______

DESCRIPTION EXAMPLE: Population Size = 20, Sample Size = 3

SAMPLE SELECTION TABLE

1	(2)	3	4	5	6	7	8	9	10	(11)	12	13	14	(15)	16	17	18	19	20	~~21~~	~~22~~	~~23~~	~~24~~	~~25~~
~~26~~	~~27~~	~~28~~	~~29~~	~~30~~	~~31~~	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	~~37~~	~~38~~	~~39~~	~~40~~	~~41~~	~~42~~	~~43~~	~~44~~	~~45~~	~~46~~	~~47~~	~~48~~	~~49~~	~~50~~
~~51~~	~~52~~	~~53~~	~~54~~	~~55~~	~~56~~	~~57~~	~~58~~	~~59~~	~~60~~	~~61~~	~~62~~	~~63~~	~~64~~	~~65~~	~~66~~	~~67~~	~~68~~	~~69~~	~~70~~	~~71~~	~~72~~	~~73~~	~~74~~	~~75~~
~~76~~	~~77~~	~~78~~	~~79~~	~~80~~	~~81~~	~~82~~	~~83~~	~~84~~	~~85~~	~~86~~	~~87~~	~~88~~	~~89~~	~~90~~	~~91~~	~~92~~	~~93~~	~~94~~	~~95~~	~~96~~	~~97~~	~~98~~	~~99~~	~~100~~

RANDOM NUMBER TABLE

43	94	16	27	33	70	20	48	59	17	74	51	26	7	56	53	77	91	5	98	36	37	10	34	29
81	21	83	71	66	4	40	30	35	52	57	79	44	63	88	28	32	86	100	97	22	95	23	76	90
99	45	80	6	62	13	14	31	58	38	82	64	75	49	25	41	92	54	11	18	93	1	96	42	87
65	19	50	24	15	8	68	2	55	39	73	12	9	67	84	61	46	78	89	72	60	47	85	3	69
68	36	58	84	(2)	34	86	30	59	90	53	94	7	100	74	19	91	48	51	63	15	77	95	6	17
23	64	32	76	4	66	40	24	45	60	12	10	46	78	3	31	99	87	50	8	37	80	11	33	73
70	83	57	52	26	47	1	54	79	56	75	72	89	69	9	85	39	5	14	49	92	21	29	88	20
13	65	18	62	97	16	96	22	82	35	38	42	93	98	28	25	61	27	41	44	55	67	81	71	43
88	49	77	9	46	69	71	64	93	84	5	50	61	83	81	52	16	26	76	21	99	95	74	85	55
22	40	70	1	98	18	100	25	39	59	32	34	94	17	37	47	90	53	(15)	51	19	24	96	97	2
29	89	65	20	78	68	33	45	27	72	31	23	44	14	79	56	87	57	54	41	35	8	80	62	6
73	86	10	91	3	75	66	58	92	60	38	11	82	42	4	67	13	43	12	48	30	7	63	36	28
72	67	20	81	23	27	100	44	3	77	49	42	22	11	41	89	54	18	15	5	69	39	13	29	79
25	58	19	63	78	43	4	30	52	73	38	85	98	84	10	14	35	46	71	7	47	28	95	55	8
33	37	26	74	90	83	75	92	1	24	96	82	65	62	60	94	61	36	91	59	76	32	21	86	64
16	97	70	88	45	99	93	12	40	34	87	66	17	56	2	80	6	50	48	51	31	53	68	57	9
12	59	94	95	3	67	32	5	27	51	49	20	26	42	76	69	2	58	60	53	86	45	92	15	74
47	36	66	22	93	79	41	57	(11)	98	28	8	34	75	21	82	83	68	48	54	73	50	37	52	78
13	46	23	4	97	56	18	40	33	31	100	16	7	90	81	84	70	39	85	1	25	62	88	77	61
71	64	44	43	91	10	38	99	87	24	63	80	29	96	19	89	17	6	72	14	65	30	9	35	55
50	88	61	57	15	5	40	64	94	97	56	59	76	74	93	36	67	84	2	82	71	48	91	83	7
99	8	27	77	60	18	37	73	28	10	43	86	63	80	66	68	19	44	31	98	47	9	25	62	70
95	34	52	96	55	81	75	24	87	29	12	23	21	16	35	85	58	1	11	14	69	51	100	42	92
65	45	32	22	72	38	6	54	3	30	78	17	20	26	46	13	90	4	33	79	41	49	39	53	89

INSTRUCTIONS

(1) IN THE SAMPLE SELECTION TABLE, CROSS OUT ALL NUMBERS THAT ARE GREATER THAN THE LOT SIZE.

(2) DIVIDE THE REMAINING NUMBERS IN THE SAMPLE SELECTION TABLE INTO AS MANY APPROXIMATELY EQUAL SUBGROUPS AS THERE ARE SAMPLES TO BE DRAWN.

(3) USING THE RANDOM NUMBER TABLE, SELECT A SINGLE SAMPLE FROM EACH OF THE SUBGROUPS.

FIGURE 6 Worksheet for selection of stratified random sample from a lot of as many as 100 items.

C_o = \$10/yd^2 (total in-place cost of overlay),
L_d = 30 years,
L_o = 10 years,
R_{inf} = 4 percent, and
R_{int} = 8 percent.

Substituting these values in Equation 1, with several possible values for expected life, produces the following results:

Expected Life, L_e (years)	*Fraction of Intended Design Life*	*Appropriate Pay Factor (%)*
35	1.17	104
30	1.00	100
25	0.83	95
20	0.67	88
15	0.50	80
10	0.33	71
5	0.17	60
0	0.0	46

Although Equation 1 was derived by summing a geometric series, a simple check calculation is possible. If the pavement in this example were to fail exactly 10 years prematurely, then all the planned future overlays remain precisely as scheduled except for the addition of one new overlay at Year 20. Using the present in-place overlay cost of \$10/yd^2 and the inflation rate of 4 percent, Equation 2 gives the cost of this added overlay at the time it is installed.

$$\text{Future cost} = \$10(1.04)^{20} = \$21.91/\text{yd}^2 \qquad (2)$$

Then, because the pay adjustment is applied at the time of initial construction, this value is converted back to present worth using the interest rate of 8 percent.

$$\text{Present worth} = \$21.91/(1.08)^{20} = \$4.70/\text{yd}^2 \qquad (3)$$

Equation 3 gives the appropriate amount of pay adjustment to be withheld. Then, using the base price for the initial pavement of C_p = \$40/yd^2, this is expressed as a percent pay factor in Equation 4. This is seen to agree exactly with the value calculated with Equation 1 for an expected life of L_e = 20 years.

$$PF = 100(40 - 4.70)/40 = 88 \qquad (4)$$

The final step in the development of the pay schedule requires a suitable relationship between quality and performance. In the case of pavement, the fatigue relationships in the AASHTO Design Guide (*16*) can be used to estimate the expected life (L_e) as a function of the quality parameter. For highway structures, or other items for which quality-performance relationships are not readily available, engineering judgment and experience must be used to estimate service life as a function of quality. At the time of this writing, there is considerable ongoing research to better establish the performance relationships necessary for this type of approach.

There are some interesting consequences of the liquidated-damages approach. Because the pay adjustments are based on the economic impact of a departure from the specified quality level, they may be positive as well as negative. For quality in excess of the design level, the transportation agency receives a tangible benefit in terms of greater performance or service life, and accordingly this method awards a small bonus. Because it has become conventional to apply pay adjustments in the form of pay factors that represent a percentage of the base price of the construction item, the pay schedule may appear to be more or less severe depending upon the actual magnitude of the base price. In reality, this procedure equates the pay adjustment directly to the estimated gain or loss experienced by the transportation agency, which, in our opinion, is a fair and equitable approach. And finally, because it is based on the well-established principle of liquidated damages, it is believed to be more defensible than some of the earlier methods. Some of the legal considerations underlying this approach are discussed further in Section 28.

11. JUSTIFICATION FOR BONUS CLAUSES

The bonus clause, sometimes referred to as an incentive provision, was supported for several years by FHWA on an experimental basis (*17*). It was eventually concluded that, in appropriate circumstances, this approach was not only cost-beneficial but decidedly in the public interest. Subsequently, all restrictions were removed (*18*) and the bonus clause has come into more common use.

The foregoing refers to bonus clauses for either early completion or superior quality. For a quality assurance program, it is the latter application that is of interest. In the case of pavement, for example, well-established fatigue relationships (*16*) clearly demonstrate that improved quality extends the expected service life. Design and construction forces have expressed the opinion that similar benefits would result from higher-quality construction of structures, either by extending the life or reducing the amount of periodic maintenance, or both. A longer service life and reduced maintenance translate directly into cost savings for the transportation agency. The bonus provision provides additional incentive to accomplish this by sharing some of the economic benefit with the contractors who are capable of achieving higher levels of quality.

Another benefit that should not be overlooked is the positive psychological effect of providing a reward for excellent performance rather than just a penalty for poor performance. But the primary justification for a bonus provision has to do with fairness to the contractor. Acceptance procedures based on percent defective (or its counterpart, percent within limits) require at least a small bonus in order to award an average pay factor of 100 percent when the contractor produces quality precisely at the level that has been defined as acceptable. The technical explanation for this has been presented in another recent publication (*15*).

12. ADVANTAGE OF CONTINUOUS PAY SCHEDULES

Although continuous (equation-type) and stepped pay schedules can be constructed to have essentially the same long-term performance as indicated by their OC curves, there is a distinct advantage associated with the continuous form. When the true quality level of the work happens to lie close to a boundary in a stepped pay schedule, the quality estimate obtained from the sample may fall on either side of the boundary, primarily because of chance. Depending upon which side of the boundary the estimate falls, there may be a substantial

difference in pay level, which may lead to disputes over measurement precision, round-off rules, and so forth. This potential problem can be completely avoided with continuous pay schedules that provide a smooth progression of payment as the quality varies.

13. DEVELOPMENT OF ACCEPTANCE PLAN

The following is a simplified example to illustrate most of the foregoing concepts. The values chosen for this example are believed to be realistic but are presented primarily to demonstrate the logical sequence of steps that must be followed in developing a statistical specification.

Since pavement life is strongly related to thickness, it is desired to develop an acceptance procedure that will provide a strong incentive to the construction industry to achieve the desired thickness. The procedure must also protect against accepting pavement that is not of sufficient thickness. Pavement thickness will be determined by taking an appropriate number of cores at random locations and measuring their length in accordance with some standard procedure.

On the basis of engineering judgment and an analysis of historical data, it has been determined that the pavement will be considered satisfactory if at least 90 percent of it is greater than the design thickness. Therefore, the AQL may be considered to be 10 percent defective and it is desired that this level of quality have a relatively high probability of acceptance or have an expected pay factor close to 100 percent, depending upon which type of acceptance procedure is used. At the other extreme, if 50 percent or more of the pavement is less than the design thickness, it has been decided that this will be regarded as rejectable (RQL) and an appropriately low probability of acceptance, or a correspondingly low expected pay factor, is desired.

Plan A

If the simplest form of acceptance procedure is desired, requiring the user to make no statistical calculations or to consult no tables, an attributes procedure would be chosen. Only the developer of the acceptance plan would be required to make the necessary calculations or use appropriate tables to verify that the procedure has a satisfactory OC curve. One such plan might require that $N = 15$ cores be taken and, provided that no more than $C = 4$ are less than the design thickness, the lot would be judged acceptable. In lieu of plotting the OC curve, the same information is presented in Table 1. It can be seen that the basic requirements have been well satisfied, since AQL work will nearly always be accepted, whereas RQL work has a very low probability of acceptance.

Plan B

If the transportation agency has confirmed that pavement thickness is at least approximately normally distributed and is willing to use slightly more sophisticated procedures, significant economies can be realized by developing a variables plan. To illustrate the savings in sample size to achieve the same result, consider a variables plan with a sample size of $N = 10$. The quality index is computed in accordance with Equation 5 and the acceptance requirement is given by Equations 6, 7, or 8.

TABLE 1 COMPARISON OF OPERATING CHARACTERISTICS OF PLANS A, B, AND C

Lot Percent Defective	Probability of Acceptance		Expected Pay Factor, Plan C (%)
	Plan A	Plan B	
10 (AQL)	0.99	0.99	99.9
20	0.84	0.82	96.9
30	0.52	0.50	91.5
40	0.22	0.21	83.2
50 (RQL)	0.06	0.06	73.1
60	0.01	0.01	63.2
70	0.0	0.0	55.6

$$Q = (\overline{X} - L)/S \tag{5}$$

where

Q = quality index,
$\overline{X}$ = sample mean,
S = sample standard deviation, and
L = specification limit (design thickness for this example).

Form 1:

$$Q \geq 0.539 \tag{6}$$

Form 2:

$$\overline{X} \geq L + 0.539S \tag{7}$$

Form 3:

$$PD \leq 30 \tag{8}$$

Form 2 does not actually require the computation of the quality index, and Form 3 requires the estimation of the lot percent defective (PD) from an appropriate table, such as that shown in Figure 7. The value of 0.539 in Equations 6 and 7 is the acceptance constant selected to produce the desired OC curve. The maximum allowable estimated percent defective of 30 in Equation 8 serves the same purpose.

For a single-limit specification, such as this example, any of these forms is equally suitable and will produce the OC curve shown in Table 1. For a double-limit specification, one having both lower and upper limits, only the third form is appropriate if a consistent OC curve is to be obtained.

Plan C

Plans A and B produce pass-or-fail decisions and do not use adjusted pay schedules. If the transportation agency desires the additional practicality of a plan that will permit the acceptance or marginally deficient work at reduced payment, it must first develop at least an approximate relationship between quality level and appropriate payment. For this example, it is assumed that the linear relationship given by

VARIABILITY-UNKNOWN PROCEDURE — SAMPLE SIZE 10 — STANDARD DEVIATION METHOD

Q	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	50.00	49.62	49.23	48.85	48.46	48.08	47.70	47.31	46.93	46.54
0.1	46.16	45.78	45.40	45.01	44.63	44.25	43.87	43.49	43.11	42.73
0.2	42.35	41.97	41.60	41.22	40.84	40.47	40.09	39.72	39.34	38.97
0.3	38.60	38.23	37.86	37.49	37.12	36.75	36.38	36.02	35.65	35.29
0.4	34.93	34.57	34.21	33.85	33.49	33.13	32.78	32.42	32.07	31.72
0.5	31.37	31.02	30.67	30.32	29.98	29.64	29.29	28.95	28.61	28.28
0.6	27.94	27.60	27.27	26.94	26.61	26.28	25.96	25.63	25.31	24.99
0.7	24.67	24.35	24.03	23.72	23.41	23.10	22.79	22.48	22.18	21.87
0.8	21.57	21.27	20.98	20.68	20.39	20.10	19.81	19.52	19.23	18.95
0.9	18.67	18.39	18.11	17.84	17.56	17.29	17.03	16.76	16.49	16.23
1.0	15.97	15.72	15.46	15.21	14.96	14.71	14.46	14.22	13.97	13.73
1.1	13.50	13.26	13.03	12.80	12.57	12.34	12.12	11.90	11.68	11.46
1.2	11.24	11.03	10.82	10.61	10.41	10.21	10.00	9.81	9.61	9.42
1.3	9.22	9.03	8.85	8.66	8.48	8.30	8.12	7.95	7.77	7.60
1.4	7.44	7.27	7.10	6.94	6.78	6.63	6.47	6.32	6.17	6.02
1.5	5.87	5.73	5.59	5.45	5.31	5.18	5.05	4.92	4.79	4.66
1.6	4.54	4.41	4.30	4.18	4.06	3.95	3.84	3.73	3.62	3.52
1.7	3.41	3.31	3.21	3.11	3.02	2.93	2.83	2.74	2.66	2.57
1.8	2.49	2.40	2.32	2.25	2.17	2.09	2.02	1.95	1.88	1.81
1.9	1.75	1.68	1.62	1.56	1.50	1.44	1.38	1.33	1.27	1.22
2.0	1.17	1.12	1.07	1.03	0.98	0.94	0.90	0.86	0.82	0.78
2.1	0.74	0.71	0.67	0.64	0.61	0.58	0.55	0.52	0.49	0.46
2.2	0.44	0.41	0.39	0.37	0.34	0.32	0.30	0.29	0.27	0.25
2.3	0.23	0.22	0.20	0.19	0.18	0.16	0.15	0.14	0.13	0.12
2.4	0.11	0.10	0.09	0.08	0.08	0.07	0.06	0.06	0.05	0.05
2.5	0.04	0.04	0.03	0.03	0.03	0.02	0.02	0.02	0.01	0.01
2.6	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00

NUMBERS IN BODY OF TABLE ARE ESTIMATES OF PERCENT DEFECTIVE CORRESPONDING TO SPECIFIC VALUES OF Q, THE QUALITY INDEX. FOR Q VALUES GREATER THAN OR EQUAL TO ZERO, THE PERCENT DEFECTIVE ESTIMATE MAY BE READ DIRECTLY FROM THE TABLE. FOR Q VALUES LESS THAN ZERO, THE TABLE VALUE MUST BE SUBTRACTED FROM 100.

FIGURE 7 Typical table for the estimation of percent defective.

Equation 9 is a suitable approximation within the region in which most of the quality estimates are likely to fall. As in Plan B, the use of variables procedures (percent defective or percent within limits) assumes that the product being inspected is at least approximately normally distributed.

$$PF = 102 - 0.2PD \tag{9}$$

where *PF* is the appropriate pay factor (percent) and *PD* is the percent defective.

Equation 9 can also be used as the pay equation for the specification. It can be seen by inspection that the AQL has been taken to be $PD = 10$ because when this value is substituted, it produces a pay factor of 100 percent. This equation also provides a small bonus for superior quality, awarding a maximum pay factor of 102 percent when the estimated percent defective is zero. At the other extreme, where the maximum value of $PD = 100$, the minimum pay factor produced by Equation 9 is 82 percent. In actual practice, most transportation agencies would want to include an RQL provision to override this process at some unacceptably large level of percent defective. In order to compute the OC curve presented in Table 1, it was assumed that a pay factor of 50 percent was assigned in lieu of requiring removal and replacement whenever the percent defective equaled or exceeded $PD = 50$.

Because of the tendency for averages to converge on the true population parameter (see Section 16), lower sample sizes can be used with pay adjustment procedures. For this example, a sample size of $N = 5$ has been selected. The acceptance procedure would require that the quality index be computed with Equation 5, the percent defective estimate (*PD*) be obtained from a table similar to that shown in Figure 7 (except designed for a sample size of $N = 5$), and the pay factor for the lot be computed with Equation 9.

The purpose of performing an analysis such as this is to ensure that the specification will be sound and defensible, that it will provide sufficient incentive to the construction industry to control the quality of the work, and that it will furnish adequate protection to the transportation agency against accepting defective work. Table 1 provides the information necessary to decide which, if any, of the plans is suitable and whether or not further refinements would be desirable.

14. ACCURACY, PRECISION, AND BIAS

Not all material test methods or statistical procedures perform equally well. In order to provide a quantitative means of comparing different procedures, statisticians have introduced the concepts of accuracy, precision, and bias.

- *Accuracy*: A procedure is said to be accurate if the mean of a distribution of measurements tends to coincide with the true mean of the population.
- *Precision*: A procedure is said to be precise if repeat tests or measurements on identical samples tend to reproduce the same value.
- *Bias*: Bias is a measure of inaccuracy, that is, the degree to which the mean of a distribution of measurements tends to be displaced from the true population value.

Perhaps the best visual impression of these terms can be obtained by imagining a marksman shooting at a target (*19*), as shown in Figure 8. The first case, in which the measurements are both accurate and precise, represents the conditions

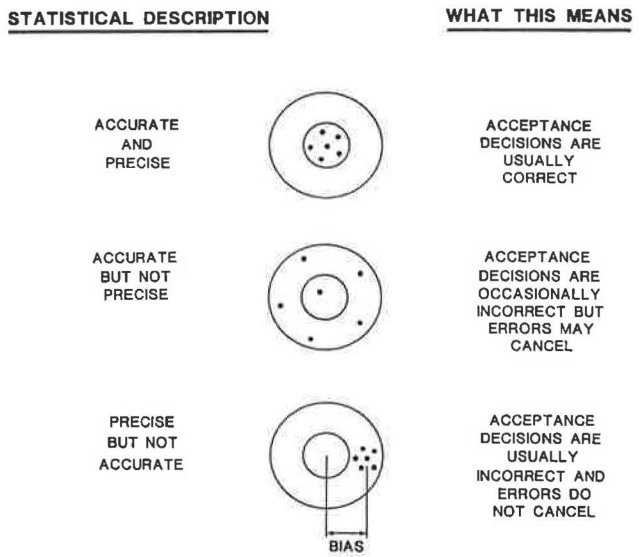

FIGURE 8 Conceptual representation of accuracy, precision, and bias.

desired for SQA specifications. Both the test methods and the statistical estimation procedures can be relied upon to provide accurate information about the quality of the work, enabling the transportation agency to make fair and appropriate acceptance decisions. The statistical acceptance procedures presented in this paper are known to be accurate, although some are more precise than others. A direct comparison of standard deviation, range, and attributes plans is presented in Part 2 and similar comparisons are made in both Sections 13 and 16 in this part of the paper.

The second example in Figure 8 shows a process that is accurate but not precise. This is less desirable, of course, but does not automatically rule out the use of such procedures. In fact, there are three ways in which the use of less precise procedures can be justified. First, if the criticality of the item being inspected is relatively low, such that the cost of making an incorrect acceptance decision is not particularly great (for either the transportation agency or the contractor), then a less precise procedure may be adequate. If the less precise procedure also happens to be simpler, faster, or less expensive, this would further justify its use. Second, if the item being inspected is a critical component that would warrant a greater degree of precision in the acceptance process, this can be accomplished with the use of a larger sample size. This, in effect, improves the precision of the overall process. Finally, if the acceptance procedure uses an adjusted pay schedule, somewhat poorer precision on individual lot determinations is not necessarily detrimental provided there are a sufficient number of lots to allow the overall pay factor for the project to average out to the value that is considered appropriate. This is discussed in more detail in Section 16.

The third example in Figure 8 is more problematical. Bias is a particularly insidious problem for two reasons: it can greatly distort the results that are obtained, and there usually is no obvious warning that it is present. Although there have been rare situations in which statisticians have been willing to tolerate a small amount of bias in exchange for a substantial improvement in precision, we doubt that such a compromise would be practical for most SQA applications. As a minimal defense against bias, we recommend using only standard and well-established test methods and statistical procedures.

A fourth case, in which the measurements are neither accurate nor precise, is not shown in Figure 8 and is of no practical use. To guard against such an undesirable situation, it is necessary to investigate the accuracy and precision of any procedures being considered for quality assurance applications.

15. SAMPLE COVERAGE AND QUALITY ASSURANCE

An idea that warrants reconsideration is the relatively common practice of relating sample size to lot size. For example, if it were customary to perform $N = 3$ tests on a lot of 100 yd^3 of concrete, some might believe that $N = 6$ tests would be required for a lot of 200 yd^3 in order to maintain the same level of quality assurance. However, as pointed out in Section 8, the level of protection is determined almost entirely by the sample size, not the lot size. Provided that the 200-yd^3 lot is produced under one set of conditions and a suitable stratified random sampling procedure is used, the $N = 3$ tests will provide the same degree of quality assurance for the larger lot. For this reason, it makes more sense to base sample sizes on criticality rather than the volume of the item being constructed. For example, a prestressed beam of 50 yd^3 might well warrant a sample size of $N = 6$ tests, whereas 300 yd^3 of concrete pavement might be suitably controlled with $N = 3$ tests. Recognizing this fact may permit a reduction in overall sampling effort or at least allow the current effort to be used more effectively.

16. AVERAGING POWER OF STATISTICS

A criticism that is sometimes leveled against statistical specifications is that, with the relatively small sample sizes typically used by transportation agencies, there may be considerable uncertainty in the decisions made on individual lots. Although this observation is theoretically correct, it is largely mitigated by the manner in which statistical acceptance procedures have been applied. It is a well-known statistical principle that sample averages tend to fall closer to the true population values as the number of items making up the average increases. The conventional method of applying pay adjustments allows this statistical principle to operate in a very desirable way. Although the pay adjustment for any individual lot may be somewhat lower or higher than it ideally should be, the overall average pay factor for the project will be very close to the appropriate value, provided the project includes a sufficient number of lots to allow the averaging process to operate. Computer simulation (discussed in Section 7) provides an effective tool to demonstrate this principle.

One word of caution must be added, however. If the acceptance procedure includes an RQL provision to require the removal and replacement of a construction item at some seriously low level of quality, the averaging process is neither operative nor appropriate. In this case, the OC curves must

be analyzed to determine if the risk of accepting truly defective work is at a suitably low level.

17. APPROPRIATE MEASURES OF VARIABILITY

The use of variables procedures (percent defective or percent within limits) is motivated by the belief that it is important to control the variability as well as the mean level of the construction characteristic in question (as discussed in Sections 3 and 4). To this end, it is important that appropriate measures of variability be used. There are two cautions to be offered—one pertaining to the statistical measure of variability and the other relating to the source of the variability.

In many of the earlier SQA specifications, the range was often used as the statistical measure of variability because it was easy to understand and compute. However, as engineers have become more familiar with statistical methods, they have begun to realize that this choice has resulted in a substantial loss of efficiency. The standard deviation, which can now be computed with a single keystroke on most modern scientific calculators, makes considerably more efficient use of the data. Table D3 in the textbook by Duncan (*3*) provides several illustrations of this. For example, the range computed from a sample of size $N = 12$ has $\nu = 9$ degrees of freedom, whereas the standard deviation computed from a sample of size $N = 10$ also has $\nu = N - 1 = 9$ degrees of freedom. Since the magnitude of the degrees of freedom provides an indication of the discriminating power of an acceptance plan, it is seen that a standard deviation plan can accomplish the same result as a range plan but with a significant savings in sampling and testing costs. To demonstrate that the two plans are equivalent, as indicated by their OC curves, tables such as those in the appendix of AASHTO Standard R9, Method A, are consulted (*20*). The appropriate values are reproduced in Table 2.

The other caution relates to the source of the variability, regardless of the parameter used to measure it. This question often arises in the assessment of the capability of the construction industry to meet a new acceptance procedure that is under development. Since statistical acceptance procedures are applied on a lot-by-lot basis, it is the typical within-lot variability plus the ability to control the mean that determine a contractor's ability to meet the specification. It would be inappropriate, for example, to gather data from an entire construction season and compute an overall standard deviation. This result would be inflated by lot-to-lot variation and would be misleadingly large. For a more realistic assessment, the standard deviations within a large number of lots should be computed and pooled in the appropriate statistical manner (*21*, p. 113) as indicated by Equation 10. To this it would be necessary to add a relatively small component to account for the variabilty of the process mean.

TABLE 2 TWO ACCEPTANCE PLANS WITH IDENTICAL OPERATING CHARACTERISTIC CURVES

	Probability of Acceptance	
Percent Defective	Standard Deviation Plan (N = 10) (M = 28)[a]	Range Plan (N = 12) (M = 28)[a]
10	0.98	0.98
20	0.77	0.77
30	0.43	0.43
40	0.17	0.17
50	0.05	0.05
60	0.01	0.01
70	0.0	0.0

[a]Maximum allowable estimated percent defective.

$$S_p = \left[\sum(N_i - 1)\, S_i^2 \Big/ \sum(N_i - 1)\right]^{1/2} \qquad (10)$$

where

S_p = pooled standard deviation,
N_i = sample sizes of individual lots, and
S_i = standard deviations of individual lots.

18. MINOR PROBLEM WITH VARIABLES PLANS

Although such occurrences are rare, it will occasionally happen when using variables acceptance procedures that a lot may be judged rejectable, or receive a pay reduction, even though none of the individual test results fall outside the specification limits. Provided that no fundamental assumptions such as a normal population or random sampling have been violated, this is a theoretically correct result. The proper interpretation is that, on the basis of the mean and standard deviation estimated from the sample, the population percent defective is unacceptably large.

In the case of a one-sided specification, this situation may also be caused by one or more outliers, test results that deviate unusually far from the norm because of some assignable cause such as equipment malfunction or operator error. The outliers could also be caused by a nonnormal population, such as would occur if two or more distinctly different normal populations were combined in the same lot. Because such a result has the appearance of being unfair, and may in fact be an indication of a breakdown in the sampling and testing process, an investigation of the cause would be warranted if it were to occur more than a very small percentage of the time.

19. PROBLEM WITH VARIABILITY-KNOWN PROCEDURES

The variables acceptance procedures described for Plan B in Section 13 are termed variability-unknown procedures because the variability of the lot is acknowledged to be unknown and is estimated by the sample standard deviation (S). Although it is possible to construct variability-known plans, in which the standard deviation is not computed but is assumed to equal some typical value for the item in question, we urge considerable caution in applying such an approach. We believe that there are a few, if any, construction items for which the variability can confidently be regarded as known. Variables approaches are usually selected when it is considered important to control the variability of the lot. Variability-known plans provide no particular incentive to do this and,

in our opinion, are not suitable for the majority of construction specification applications.

20. DUAL ACCEPTANCE PROCEDURES

Some specifications in current use contain dual acceptance criteria. In addition to the primary requirement applied to some parameter computed from the entire sample (mean, percent defective, etc.), there may be a secondary requirement applied to the individual values. A dual acceptance procedure in common use requires the mean of the sample to exceed some limit and each individual value of the sample to exceed some lower limit. Often, the primary requirement is intended to dominate, and the secondary requirement comes into play less frequently.

The probabilities of passing the two requirements separately are readily calculated by conventional statistical techniques. However, although it may be tempting to assume that the probability of passing both requirements is simply the product of these two separate probabilities, this is not the case. Because the chance occurrence of unusually low or high test values will have a similar effect on the likelihood of passing either requirement, the two probabilities are not independent but are positively correlated to some unknown degree. This lack of independence requires that a different approach be taken.

Although it is not possible to compute the desired probability directly, it can be approximated by computing lower and upper bounds. This furnishes interval estimates that, in most cases, are sufficiently narrow to be of practical use. Figure 9 shows typical bounding OC curves obtained by this method. The theoretical development may be found in either of two recent publications (*8,22*), in which the following expression is derived:

$$P_1P_2 \leq P(\text{Accept}) \leq \text{Min}(P_1, P_2) \quad (11)$$

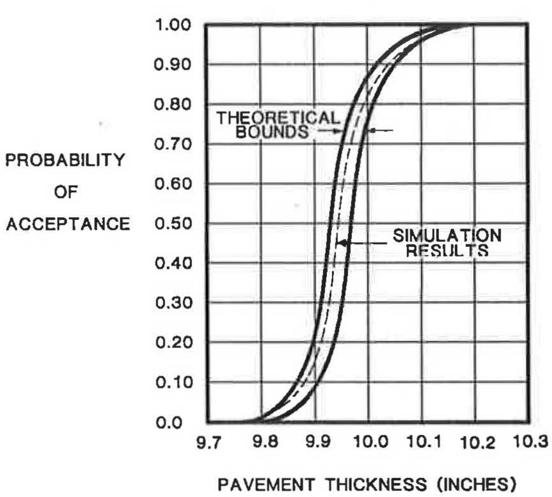

FIGURE 9 OC curves for a dual acceptance procedure.

where

P_1 = probability of passing the first requirement,
P_2 = probability of passing the second requirement, and
$P(\text{Accept})$ = overall probability of acceptance.

21. RETESTING PROVISIONS

Retesting provisions have generated an inordinate amount of confusion. This may be due to the different reasons for their use, the different ways in which the retest results can be processed, and the advantages and disadvantages associated with each.

There are at least four reasons why a retesting provision might be made an integral part of an acceptance plan:

1. To confirm that the work is truly defective before imposing a severe consequence, such as requiring removal and replacement or assigning a minimum pay factor.
2. To produce a more desirable OC curve that properly balances the risks between the transportation agency and the contractor.
3. To make more efficient use of limited sampling and testing resources. A reduced sampling effort may be sufficient to identify work that is clearly good or clearly defective. When the work falls in the marginal area between these two extremes, the additional information provided by the retest enables the transportation agency to make an appropriate acceptance decision.
4. To guard against a possible breakdown in any of the steps of the sampling and testing process.

If a retest provision is to be used, it must be described explicitly in the contract documents, including precisely how the retest results are to be processed. There are two distinctly different ways to do this: they may be combined with the original test results or they may be used in place of the original test results.

An advantage of the first method is that it makes maximum use of all the available information. Advocates of this method argue that there is a cost associated with each sample and that it is wasteful to discard any valid information. An opposing viewpoint would question whether the original sample was truly valid. If the low quality level is the result of some malfunction of the testing process, then it would be more appropriate to disregard the contaminated data.

This is a question of philosophy that each agency must answer for itself. However, if the decision is made to combine the retest results with those obtained from the initial sample, caution must be exercised in computing the OC curve for this procedure. Since the probabilities of failing the original test and passing the retest are correlated to some degree, the overall probability must be determined either by the boundary method (Section 20) or by computer simulation (Section 7). If the first and second series of tests are designated by the letters A and B, Equation 12 gives the bounds for the overall probability of acceptance by this procedure (*8,22*).

$$\text{Max}(P_A, P_B) \leq P(\text{Accept}) \leq P_A + (1 - P_A)P_B \quad (12)$$

where

$$P_A = \text{probability of acceptance at Stage A,}$$
$$P_B = \text{probability of acceptance at Stage B, and}$$
$$P(\text{Accept}) = \text{overall probability of acceptance.}$$

22. ZERO PAY FACTORS

Over the years, as statistical acceptance procedures have evolved, the argument has occasionally been heard that a zero pay factor is inappropriate. The position has been stated that, when a transportation agency chooses, for practical reasons, not to require the removal and replacement of an extremely defective item, this is an acknowledgment that the item can provide some limited degree of service and that therefore it has some minimum value greater than zero.

We offer a counterargument. Although the approach that we advocate for quantifying pay adjustments (Section 10) has not yet resulted in a pay schedule with a zero pay factor, we see no theoretical reason why this would be inappropriate. There will occasionally be situations in which it is judged more practical to accept a limited amount of performance than to undertake a major repair at the time of construction. However, if a major repair were eventually required and were a direct consequence of inferior workmanship, and the extra expense that must be borne by the transportation agency were to equal or exceed the initial cost of the construction item, then we believe that the legal concept of liquidated damages would support a pay factor of zero or less (see Section 28).

23. AVAILABILITY OF STATISTICAL TABLES

If SQA methods are to be promoted, they must be made as easy to understand and implement as possible. Unfortunately, some of the earlier attempts were hindered by the necessity of using statistical tables that were not in the most useful form for transportation applications. Other tables that would have been extremely useful simply did not exist. Several new tables have recently been developed (*4,20*) that greatly simplify the construction of OC curves for both attributes and variables acceptance plans. A still more extensive series of tables is currently in preparation (*23*). Computer programs have been developed that are capable of generating a wide variety of useful tables, some examples of which are given in Figures 10–12 as well as in Figure 7.

24. CLASSICAL VERSUS BAYESIAN METHODS

One of the oldest debates in the field of statistics concerns the relative merits of classical versus Bayesian procedures. Some statisticians choose to align themselves in one camp or the other, whereas others seem to believe that both methods have their appropriate uses. We wish to offer an argument for this latter point of view as it applies to quality control and quality assurance.

The term "quality control" is generally interpreted to refer to those actions taken by the contractor or producer to maintain the necessary quality of production to meet the requirements of the specification. The term "quality assurance" may refer to the entire system involving both quality control and acceptance testing but usually is associated more strongly with the acceptance testing activities of the transportation agency.

In judging the quality of a lot, for example, Bayesian procedures would use not only the results from that particular lot but also data from previous lots, combined in an appropriate manner (*24*). Proponents of this method argue that it is both logical and useful to make formal use of existing prior knowledge (not necessarily data alone) in order to improve the likelihood of making a correct decision. Classical statisticians, on the other hand, would argue that the use of any information other than that from the particular lot in question has the potential for biasing the decision on that lot and that, in the long run, fewer correct decisions would be made.

We believe that in order for prior knowledge to be useful, the assumption must be made that the process has remained essentially unchanged with time, that is, that there has been no actual drift in the process mean or standard deviation and that any fluctuations in the quality of the output are simply the result of inherent random variation. The only problem with this approach, at least as far as acceptance testing is concerned, is that it assumes away the very problem that acceptance procedures are designed to guard against—a true shift in quality level, whether accidental or intentional. Consequently, we favor the classical approach to acceptance testing.

However, we think the Bayesian approach might be superior in the area of process control, that portion of the overall quality assurance program that is traditionally the responsibility of the contractor. To keep the process under control, the contractor must continually react to information from the field and make timely adjustments as necessary. Because of the random variation of individual lots, however, no single lot provides a very reliable indicator of the need for a process adjustment. The Bayesian procedure, because it integrates current lot information with previous production data, may provide a more reliable basis upon which process control decisions could be made. This application seems particularly appropriate because it is the contractor who has both ready access to process data and the motivation to control the process effectively. Perhaps the academic community, whose consulting services are frequently engaged by the construction industry, could develop a simple and practical way to accomplish this.

25. MOVING AVERAGES

More generally, the concept of moving averages could be referred to as a "moving sample," because its use is not limited just to the average of the test values. Although it has a certain appeal, the use of this approach as the basis for an acceptance procedure is of questionable value.

Because small sample sizes often do not provide the desired reliability for individual acceptance decisions, especially for pass-or-fail applications that do not benefit from the averaging process discussed in Section 16, it may be desirable to combine current lots with previous lots to produce a larger total sample size. This will usually be satisfactory when it is accomplished

ATTRIBUTES ACCEPTANCE PLANS LOT SIZE = INFINITE

SAMPLE SIZE (n)	ACCEPTANCE NUMBER (c)	PROBABILITY OF ACCEPTANCE FOR SELECTED LEVELS OF LOT PERCENT DEFECTIVE													
		5	10	15	20	25	30	35	40	45	50	55	60	65	70
1	0	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	0.45	0.40	0.35	0.30
2	0	0.90	0.81	0.72	0.64	0.56	0.49	0.42	0.36	0.30	0.25	0.20	0.16	0.12	0.09
3	0	0.86	0.73	0.61	0.51	0.42	0.34	0.27	0.22	0.17	0.13	0.09	0.06	0.04	0.03
3	1	0.99	0.97	0.94	0.90	0.84	0.78	0.72	0.65	0.57	0.50	0.43	0.35	0.28	0.22
4	0	0.81	0.66	0.52	0.41	0.32	0.24	0.18	0.13	0.09	0.06	0.04	0.03	0.02	0.01
4	1	0.99	0.95	0.89	0.82	0.74	0.65	0.56	0.48	0.39	0.31	0.24	0.18	0.13	0.08
5	1	0.98	0.92	0.84	0.74	0.63	0.53	0.43	0.34	0.26	0.19	0.13	0.09	0.05	0.03
5	2	1.00	0.99	0.97	0.94	0.90	0.84	0.76	0.68	0.59	0.50	0.41	0.32	0.24	0.16
6	1	0.97	0.89	0.78	0.66	0.53	0.42	0.32	0.23	0.16	0.11	0.07	0.04	0.02	0.01
6	2	1.00	0.98	0.95	0.90	0.83	0.74	0.65	0.54	0.44	0.34	0.26	0.18	0.12	0.07
7	1	0.96	0.85	0.72	0.58	0.44	0.33	0.23	0.16	0.10	0.06	0.04	0.02	0.01	0.00
7	2	1.00	0.97	0.93	0.85	0.76	0.65	0.53	0.42	0.32	0.23	0.15	0.10	0.06	0.03
8	1	0 74	0.81	0.66	0.50	0.37	0.26	0.17	0.11	0.06	0.04	0.02	0.01	0.00	0.00
8	2	0.99	0.96	0.89	0.80	0.68	0.55	0.43	0.32	0.22	0.14	0.09	0.05	0.03	0.01
8	3	1.00	0.99	0.98	0.94	0.89	0.81	0.71	0.59	0.48	0.36	0.26	0.17	0.11	0.06
9	1	0.93	0.77	0.60	0.44	0.30	0.20	0.12	0.07	0.04	0.02	0.01	0.00	0.00	0.00
9	2	0.99	0.95	0.86	0.74	0.60	0.46	0.34	0.23	0.15	0.09	0.05	0.03	0.01	0.00
9	3	1.00	0.99	0.97	0.91	0.83	0.73	0.61	0.48	0.36	0.25	0.17	0.10	0.05	0.03
10	1	0.91	0.74	0.54	0.38	0.24	0.15	0.09	0.05	0.02	0.01	0.00	0.00	0.00	0.00
10	2	0.99	0.93	0.82	0.68	0.53	0.38	0.26	0.17	0.10	0.05	0.03	0.01	0.00	0.00
10	3	1.00	0.99	0.95	0.88	0.78	0.65	0.51	0.38	0.27	0.17	0.10	0.05	0.03	0.01
15	2	0.96	0.82	0.60	0.40	0.24	0.13	0.06	0.03	0.01	0.00	0.00	0.00	0.00	0.00
15	3	0.99	0.94	0.82	0.65	0.46	0.30	0.17	0.09	0.04	0.02	0.01	0.00	0.00	0.00
15	4	1.00	0.99	0.94	0.84	0.69	0.52	0.35	0.22	0.12	0.06	0.03	0.01	0.00	0.00
15	5	1.00	1.00	0.98	0.94	0.85	0.72	0.56	0.40	0.26	0.15	0.08	0.03	0.01	0.00
20	2	0.92	0.68	0.40	0.21	0.09	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
20	3	0.98	0.87	0.65	0.41	0.23	0.11	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00
20	4	1.00	0.96	0.83	0.63	0.41	0.24	0.12	0.05	0.02	0.01	0.00	0.00	0.00	0.00
20	5	1.00	0.99	0.93	0.80	0.62	0.42	0.25	0.13	0.06	0.02	0.01	0.00	0.00	0.00
20	6	1.00	1.00	0.98	0.91	0.79	0.61	0.42	0.25	0.13	0.06	0.02	0.01	0.00	0.00
30	3	0.94	0.65	0.32	0.12	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	4	0.98	0.82	0.52	0.26	0.10	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	5	1.00	0.93	0.71	0.43	0.20	0.08	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00
30	6	1.00	0.97	0.85	0.61	0.35	0.16	0.06	0.02	0.00	0.00	0.00	0.00	0.00	0.00
30	7	1.00	0.99	0.93	0.76	0.51	0.28	0.12	0.04	0.01	0.00	0.00	0.00	0.00	0.00
30	8	1.00	1.00	0.97	0.87	0.67	0.43	0.22	0.09	0.03	0.01	0.00	0.00	0.00	0.00
50	5	0.96	0.62	0.22	0.05	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	6	0.99	0.77	0.36	0.10	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	7	1.00	0.88	0.52	0.19	0.05	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	8	1.00	0.94	0.67	0.31	0.09	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	9	1.00	0.98	0.79	0.44	0.16	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	10	1.00	0.99	0.88	0.58	0.26	0.08	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	11	1.00	1.00	0.94	0.71	0.38	0.14	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00
100	8	0.94	0.32	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	9	0.97	0.45	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	10	0.99	0.58	0.10	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	11	1.00	0.70	0.16	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	12	1.00	0.80	0.25	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	13	1.00	0.88	0.35	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	14	1.00	0.93	0.46	0.08	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	15	1.00	0.96	0.57	0.13	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	16	1.00	0.98	0.67	0.19	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	17	1.00	0.99	0.76	0.27	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	18	1.00	1.00	0.84	0.36	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

PROBABILITY OF ACCEPTANCE IS A FUNCTION OF LOT SIZE FOR ATTRIBUTES PLANS USING DISCRETE DATA. FOR VARIABLE LOT SIZES, IT WILL BE NECESSARY TO PLOT BOUNDING OPERATING CHARACTERISTIC CURVES. THE VALUES IN THIS TABLE ARE APPROPRIATE FOR BOTH SINGLE-LIMIT AND DOUBLE-LIMIT APPLICATIONS AND ARE UNINFLUENCED BY THE DISTRIBUTIONAL FORM OF THE POPULATION.

FIGURE 10 OC table for attributes acceptance plans with an infinite (nondiscrete) lot size.

by defining new, larger lot sizes and evaluating them independently (see Section 8). It may not be satisfactory, however, when it is accomplished by evaluating the work on the basis of the most recent series of tests of the desired sample size. By this procedure, if the desired sample size were $N = 5$, then Tests 1–5 would be evaluated as an acceptance lot, Tests 2–6 would make up a second lot, and so forth.

There are two problems with this latter approach. First, because each sublot appears in several different acceptance lots, there is a lack of statistical independence and the ordinary analytical tools do not apply. To date, the only way to develop the OC curve for such a procedure is by computer simulation. Second, on the basis of computer simulation tests by one of us, the use of moving averages provides no real gain in discriminating power. Because each sublot is subject to the "double jeopardy" of having several opportunities to be rejected, somewhat broader acceptance limits must be used. The net result is an OC curve almost identical to that obtained for independent lots using the same overall sampling rate.

Like the Bayesian approach discussed in Section 24, we believe that the moving average is better suited as a process control device. Because it tends to damp out statistical vari-

VARIABLES ACCEPTANCE PLANS *** VARIABILITY-UNKNOWN PROCEDURE *** STANDARD DEVIATION METHOD

SAMPLE SIZE (n)	MAXIMUM ALLOWABLE ESTIMATED PERCENT DEFECTIVE (M)	MINIMUM ALLOWABLE QUALITY INDEX (k)	PROBABILITY OF ACCEPTANCE FOR SELECTED LEVELS OF LOT PERCENT DEFECTIVE						
			10	20	30	40	50	60	70
3	34	0.556	0.89	0.71	0.52	0.35	0.22	0.12	0.05
3	36	0.492	0.91	0.74	0.56	0.39	0.24	0.13	0.06
3	38	0.425	0.93	0.78	0.60	0.42	0.27	0.15	0.07
3	40	0.357	0.95	0.81	0.64	0.46	0.30	0.17	0.08
3	42	0.287	0.96	0.84	0.68	0.50	0.33	0.20	0.09
3	44	0.216	0.97	0.87	0.71	0.54	0.37	0.22	0.11
3	46	0.145	0.98	0.89	0.75	0.58	0.41	0.26	0.13
3	48	0.073	0.98	0.91	0.79	0.63	0.46	0.29	0.15
4	28	0.660	0.88	0.66	0.44	0.27	0.14	0.06	0.02
4	30	0.600	0.91	0.70	0.48	0.29	0.16	0.07	0.02
4	32	0.540	0.93	0.74	0.52	0.33	0.18	0.08	0.03
4	34	0.480	0.94	0.77	0.56	0.36	0.20	0.10	0.03
4	36	0.420	0.96	0.81	0.60	0.40	0.23	0.11	0.04
4	38	0.360	0.97	0.84	0.64	0.44	0.26	0.13	0.05
4	40	0.300	0.97	0.86	0.69	0.48	0.30	0.15	0.06
4	42	0.240	0.98	0.89	0.72	0.53	0.33	0.18	0.07
4	44	0.180	0.99	0.91	0.76	0.57	0.37	0.20	0.09
4	46	0.120	0.99	0.93	0.79	0.61	0.41	0.23	0.10
5	26	0.692	0.89	0.65	0.40	0.22	0.10	0.04	0.01
5	28	0.632	0.92	0.69	0.45	0.25	0.12	0.04	0.01
5	30	0,572	0.94	0.73	0.49	0.28	0.14	0.05	0.01
5	32	0.513	0.95	0.77	0.54	0.32	0.16	0.06	0.02
5	34	0.455	0.96	0.81	0.58	0.36	0.18	0.07	0.02
5	36	0.397	0.97	0.84	0.63	0.40	0.21	0.09	0.03
5	38	0.339	0.98	0.87	0.67	0.44	0.25	0.11	0.03
5	40	0.282	0.99	0.90	0.71	0.49	0.28	0.13	0.04
5	42	0.225	0.99	0.92	0.75	0.54	0.32	0.15	0.05
5	44	0.169	0.99	0.93	0.79	0.58	0.36	0.18	0.06
6	24	0.740	0.89	0.62	0.35	0.17	0.06	0.02	0.00
6	26	0.678	0.92	0.67	0.40	0.20	0.08	0.02	0.00
6	28	0.618	0.94	0.71	0.45	0.23	0.10	0.03	0.01
6	30	0.558	0.95	0.76	0.49	0.27	0.11	0.04	0.01
6	32	0.500	0.97	0.80	0.55	0.31	0.14	0.05	0.01
6	34	0.442	0.98	0.84	0.60	0.35	0.16	0.06	0.01
6	36	0.386	0.98	0.87	0.64	0.39	0.19	0.07	0.02
6	38	0.329	0.99	0.90	0.69	0.44	0.23	0.09	0.02
6	40	0.274	0.99	0.92	0.74	0.49	0.27	0.11	0.03
6	42	0.219	1.00	0.94	0.78	0.54	0.31	0.13	0.04
7	22	0.796	0.88	0.57	0.29	0.12	0.04	0.01	0.00
7	24	0.732	0.91	0.63	0.34	0.15	0.05	0.01	0.00
7	26	0.670	0.93	0.68	0.39	0.18	0.06	0.02	0.00
7	28	0.610	0.95	0.73	0.44	0.21	0.08	0.02	0.00
7	30	0.550	0.97	0.78	0.50	0.25	0.10	0.03	0.00
7	32	0.492	0.98	0.82	0.55	0.29	0.12	0.04	0.01
7	34	0.435	0.98	0.86	0.61	0.34	0.15	0.05	0.01
7	36	0.379	0.99	0.89	0.66	0.39	0.18	0.06	0.01
7	38	0.324	0.99	0.91	0.71	0.44	0.21	0.07	0.02
7	40	0.269	1.00	0.93	0.75	0.49	0.25	0.09	0.02
8	22	0.792	0.90	0.58	0.28	0.11	0.03	0.01	0.00
8	24	0.727	0.92	0.64	0.33	0.13	0.04	0.01	0.00
8	26	0.665	0.95	0.70	0.38	0.16	0.05	0.01	0.00
8	28	0.604	0.96	0.75	0.44	0.20	0.07	0.01	0.00
8	30	0.545	0.98	0.80	0.50	0.24	0.08	0.02	0.00
8	32	0.488	0.98	0.84	0.56	0.28	0.11	0.03	0.00
8	34	0.431	0.99	0.87	0.62	0.33	0.13	0.04	0.01
8	36	0.375	0.99	0.90	0.67	0.38	0.16	0.05	0.01
8	38	0.320	1.00	0.93	0.72	0.44	0.20	0.06	0.01
9	20	0.855	0.87	0.52	0.22	0.07	0.02	0.00	0.00
9	22	0.788	0.91	0.58	0.27	0.09	0.02	0.00	0.00
9	24	0.724	0.94	0.65	0.32	0.12	0.03	0.01	0.00
9	26	0.661	0.96	0.71	0.38	0.15	0.04	0.01	0.00
9	28	0.601	0.97	0.76	0.44	0.18	0.05	0.01	0.00
9	30	0.542	0.98	0.81	0.50	0.22	0.07	0.01	0.00
9	32	0.484	0.99	0.85	0.56	0.27	0.09	0.02	0.00
9	34	0.428	0.99	0.89	0.62	0.32	0.12	0.03	0.00
9	36	0.373	1.00	0.92	0.68	0.38	0.15	0.04	0.01
10	20	0.853	0.89	0.51	0.21	0.06	0.01	0.00	0.00
10	22	0.786	0.92	0.59	0.26	0.08	0.02	0.00	0.00
10	24	0.721	0.95	0.65	0.31	0.10	0.02	0.00	0.00
10	26	0.659	0.97	0.72	0.37	0.13	0.03	0.01	0.00
10	28	0.598	0.98	0.77	0.43	0.17	0.05	0.01	0.00
10	30	0.539	0.99	0.82	0.50	0.21	0.06	0.01	0.00
10	32	0.482	0.99	0.87	0.57	0.26	0.08	0.02	0.00
10	34	0.426	1.00	0.90	0.63	0.31	0.11	0.02	0.00

THE ACCEPTANCE PROBABILITIES IN THIS TABLE ARE ACCURATE FOR SINGLE-LIMIT PLANS AND ARE APPROXIMATELY CORRECT FOR DOUBLE-LIMIT PLANS. FOR SINGLE-LIMIT PLANS, EITHER THE MAXIMUM ALLOWABLE ESTIMATED PERCENT DEFECTIVE (M) OR THE MINIMUM ALLOWABLE QUALITY INDEX (k) MAY BE SPECIFIED. FOR DOUBLE-LIMIT PLANS, ONLY THE MAXIMUM ALLOWABLE ESTIMATED PERCENT DEFECTIVE SHOULD BE USED.

FIGURE 11 OC table for variables acceptance plans.

ability and make long-term trends easier to discern, it has the ability to provide extremely useful guidance to contractors who must continually monitor their processes to stay within specification limits.

26. UNBALANCED BIDS

A potential problem that can affect the proper application of an adjusted pay schedule is an unbalanced bid. There are various reasons, often related to cash flow, for which contractors choose to bid unusually low on some items and unusually high on others. At one time, it was NJDOT practice to reject bids that were obviously unbalanced. In recent years, it has come to be regarded as in the public interest to accept such a bid if it is otherwise valid and is not likely to cause major cost increases due to subsequent change orders. Because it would be possible for a contractor to virtually nullify the effect of an adjusted pay schedule by underbidding on those items to which the pay schedule applies, NJDOT has

VARIABLES ACCEPTANCE PLANS *** VARIABILITY-UNKNOWN PROCEDURE *** STANDARD DEVIATION METHOD

SAMPLE SIZE (n)	MAXIMUM ALLOWABLE ESTIMATED PERCENT DEFECTIVE (M)	MINIMUM ALLOWABLE QUALITY INDEX (k)	LOT PERCENT DEFECTIVE VALUES PRODUCING THE LISTED ACCEPTANCE PROBABILITIES								
			0.99	0.95	0.90	0.80	0.50	0.20	0.10	0.05	0.01
3	30	0.679	1	4	7	12	28	48	59	68	81
3	35	0.524	3	7	10	16	32	53	63	71	84
3	40	0.357	4	10	14	21	38	58	67	75	86
3	45	0.181	6	13	18	26	44	63	72	79	88
4	25	0.750	2	5	8	12	25	42	52	60	74
4	30	0.600	3	7	10	15	29	47	56	64	76
4	35	0.450	5	10	14	19	34	51	60	68	79
4	40	0.300	7	13	17	24	39	56	65	71	82
4	45	0.150	9	17	22	29	44	61	69	76	85
5	25	0.723	3	6	9	13	25	40	49	56	69
5	30	0.572	5	9	12	17	30	45	53	60	72
5	35	0.426	7	12	16	21	34	50	58	65	76
5	40	0.282	9	15	20	26	40	55	63	69	79
5	45	0.141	12	19	24	30	45	60	67	73	82
6	20	0.869	2	5	7	11	21	34	42	49	61
6	25	0.709	4	7	10	14	25	39	47	53	65
6	30	0.558	6	10	14	18	30	44	51	58	69
6	35	0.414	8	14	17	23	35	49	56	62	73
6	40	0.274	11	17	21	27	40	54	61	66	76
7	20	0.862	3	6	8	11	21	33	40	46	58
7	25	0.701	5	8	11	15	25	38	45	51	62
7	30	0.550	7	11	15	19	30	43	50	56	66
7	35	0.407	9	15	18	23	35	48	55	60	70
7	40	0.269	12	19	23	28	40	53	59	65	74
8	20	0.858	3	6	9	12	20	32	38	44	55
8	25	0.696	5	9	12	16	25	37	44	49	60
8	30	0.545	8	12	15	20	30	42	48	54	64
8	35	0.403	11	16	19	24	35	47	53	59	68
9	20	0.855	4	7	9	12	20	31	37	43	53
9	25	0.692	6	10	12	16	25	36	42	48	58
9	30	0.542	9	13	16	20	30	41	47	53	62
9	35	0.400	12	17	20	25	35	46	52	57	66
10	20	0.853	4	7	10	13	20	30	36	41	51
10	25	0.690	7	10	13	17	25	35	41	46	56
10	30	0.539	9	14	17	21	30	41	46	51	60
10	35	0.398	12	18	21	25	35	46	51	56	65
15	15	1.037	4	6	8	10	15	23	27	31	39
15	20	0.848	6	9	11	14	20	28	33	37	45
15	25	0.683	9	13	15	18	25	33	38	42	50
15	30	0.533	12	16	19	22	30	39	43	47	55
20	15	1.036	5	7	8	10	15	22	25	29	35
20	20	0.846	7	10	12	15	20	27	31	34	41
20	25	0.680	10	14	16	19	25	32	36	40	46
30	15	1.036	6	8	9	11	15	20	23	26	31
30	20	0.844	9	12	13	15	20	26	29	31	37
30	25	0.678	13	16	18	20	25	31	34	37	42
50	10	1.277	4	6	6	8	10	13	15	17	20
50	15	1.036	7	9	10	12	15	19	21	23	27
50	20	0.843	11	13	15	16	20	24	27	29	32
100	10	1.279	5	7	7	8	10	12	14	15	17
100	15	1.036	9	11	12	13	15	18	19	20	23
100	20	0.842	13	15	16	17	20	23	25	26	29

THE ACCEPTANCE PROBABILITIES IN THE HEADING OF THIS TABLE ARE ACCURATE FOR SINGLE-LIMIT PLANS AND ARE APPROXIMATELY CORRECT FOR DOUBLE-LIMIT PLANS. FOR SINGLE LIMIT APPLICATIONS, EITHER THE MAXIMUM ALLOWABLE ESTIMATED PERCENT DEFECTIVE (M) OR THE MINIMUM ALLOWABLE QUALITY INDEX (k) MAY BE SPECIFIED. FOR DOUBLE-LIMIT APPLICATIONS, ONLY THE MAXIMUM ALLOWABLE ESTIMATED PERCENT DEFECTIVE SHOULD BE USED.

FIGURE 12 Alternative format for OC table for variables acceptance plans.

defined a "base unit price" for some pay adjustment items. In essence, this is a weighted unit price based on recent construction cost information. The contractor is paid in accordance with the bid price, but any pay adjustments are computed using the base unit price.

27. NON-PAY-ADJUSTMENT ITEMS

In any area of construction, some items are more critical than others and warrant more discerning acceptance procedures. For the NJDOT PCC specification, for example, bridge decks, columns, prestressed beams, and occasional other items are considered to be the most critical, and accordingly, larger sample sizes are required and acceptance is by pay adjustment. For each project, the items to be accepted in this manner are specifically listed in the contract documents. The remaining items are referred to as non-pay-adjustment items and are subject to less stringent acceptance requirements that lead to a pass-or-fail decision.

Recently, a minor refinement was made that provides a very practical way to deal with the occasional failure of a non-pay-adjustment item. If a non-pay-adjustment item is rejected, NJDOT has the option to regard it as a pay-adjustment item and to reevaluate it with cores at the higher sampling rate prescribed for the more critical items. This provision is seldom necessary, but it provides a formal procedure for handling such rejections if and when they occur.

28. LEGAL CONSIDERATIONS

The legal concept of liquidated damages was mentioned in connection with the development of adjusted pay schedules in Section 10. This concept may properly be applied whenever it is impossible or impractical to quantify the actual damages. Because pay schedules are designed to recoup future losses that can only be estimated, they would appear to clearly qualify for this approach.

As a general rule, liquidated-damages clauses are considered acceptable, whereas penalty clauses are not. According to Sweet,

> If enforcement of a clause would punish a breaching party by awarding an amount disproportionately high to anticipated or actual damages, the clause will not be enforced even if labeled as a damage liquidation clause. Conversely, a clause labeled a penalty will be enforced if it otherwise meets the test of damage liquidation. (*25*, p. 404)

In other words, no matter how a transportation agency chooses to label a pay adjustment clause, the magnitude of the adjustment must be reasonably commensurate with the amount of damage actually suffered. This stresses the importance of developing the necessary quality-performance relationships, as discussed in Sections 2 and 10. However, this need not be interpreted to mean that the amount of damage must be estimated with great precision. Sweet goes on to cite a Supreme Court decision that includes the following commentary on liquidated-damages clauses:

> When that intention is clearly ascertainable from the writing, effect will be given to the provision, as freely as to any other, where damages are uncertain in nature or amount or are difficult of ascertainment or where the amount stipulated for is not so extravagant, or disproportionate to the amount of property loss, as to show that compensation was not the object aimed at or as to imply fraud, mistake, circumvention or oppression. There is no sound reason why persons competent and free to contract may not agree upon this subject as fully as upon any other, or why their agreement, when fairly and understandably entered into with a view to just compensation for the anticipated loss, should not be enforced. (*25*, pp. 403–404)

In simpler language, what this appears to say is that two contracting parties may agree on the amount to be withheld in the event of noncompliance, and the courts will uphold this agreement provided that the stipulated amount is reasonably appropriate for the damages actually suffered and there is no element of deception, either consciously or inadvertently. This rationale provides a solid basis for the pay-adjustment concept in general and for the liquidated-damages approach described in Section 10 in particular.

SUMMARY AND PREVIEW

A series of fundamental concepts was presented in the belief that reliable quality assurance technology must be based on sound scientific, mathematical, and legal principles. These concepts—proven by actual field application over a period of approximately 20 years—provide the basic building blocks with which practical and effective quality assurance programs can be constructed. The guidance they provide will be useful both to individual agencies and to a task force contemplating a national policy on transportation quality assurance.

We invite a rigorous scrutiny of these concepts to accomplish either of two objectives: to confirm their validity or to improve upon them. The resulting body of knowledge can then form the technical core for an effective national policy. A series of steps that will aid in achieving this goal is developed in Part 4.

REFERENCES (Part 3)

1. *Recommended Practice for Evaluation of Strength Test Results of Concrete*. ACI Standard 214-77. American Concrete Institute, Detroit, Mich., 1977.
2. *Standard Specifications for Ready-Mixed Concrete, C94-78a*. Annual ASTM Standards, Part 14. ASTM, Philadelphia, Pa., 1978.
3. A. J. Duncan. *Quality Control and Industrial Statistics*, 3rd ed. Richard D. Irwin, Inc., Homewood, Ill., 1965.
4. R. M. Weed. Revision of a Flawed Acceptance Standard. In *Transportation Research Record 1056*, TRB, National Research Council, Washington, D.C., 1986, pp. 21–35.
5. R. T. Barros, R. M. Weed, and J. H. Willenbrock. Software Package for the Design and Analysis of Acceptance Procedures Based on Percent Defective. In *Transportation Research Record 924*, TRB, National Research Council, Washington, D.C., 1983, pp. 85–93.
6. R. M. Weed. *An Introduction to Computer Simulation*. Demonstration Projects Division, FHWA, U.S. Department of Transportation, 1976.
7. R. M. Weed. Conversational Programming: Making Computer Technology More Accessible. In *Transportation Research Record 745*, TRB, National Research Council, Washington, D.C., 1980, pp. 6–10.
8. R. M. Weed. *Statistical Specification Development*, 2nd ed. HPR Study 7788. New Jersey Department of Transportation, Trenton, 1989.
9. R. M. Weed. Computer-Assisted Random Sampling. In *Transportation Research Record 1034*, TRB, National Research Council, Washington, D.C., 1985, pp. 140–152.
10. *NCHRP Synthesis of Highway Practice 38: Statistically Oriented End-Result Specifications*. TRB, National Research Council, Washington, D.C., 1976, pp. 8–9.
11. J. H. Willenbrock and P. A. Kopac. Development of Price-Adjustment Systems for Statistically Based Highway Construction Specifications. In *Transportation Research Record 652*, TRB, National Research Council, Washington, D.C., 1977, pp. 52–58.
12. R. M. Weed. Equitable Graduated Pay Schedules: An Economic Approach. In *Transportation Research Record 691*, TRB, National Research Council, Washington, D.C., 1978, pp. 27–29.
13. R. M. Weed. Method to Establish Pay Schedules for Rigid Pavement. In *Transportation Research Record 885*, TRB, National Research Council, Washington, D.C., 1982, pp. 18–24.
14. P. Puangchit, R. G. Hicks, J. E. Wilson, and C. A. Bell. Development of Rational Pay Adjustment Factors for Asphalt Concrete. In *Transportation Research Record 911*, TRB, National Research Council, Washington, D.C., 1983, pp. 70–79.
15. R. M. Weed. Adjusted Pay Schedules: New Concepts and Provisions. In *Transportation Research Record 986*, TRB, National Research Council, Washington, D.C., 1984, pp. 32–38.
16. *Interim Guide for Design of Pavement Structures*. American Association of State Highway and Transportation Officials, Washington, D.C., 1981.
17. *Use of Incentive and Disincentive Provisions Regarding Quality and Completion Time for Federal Aid Highway Construction*.

NEEP Project 24, Notice N 5080.66. FHWA, U.S. Department of Transportation, 1977.
18. *Federal Register*. Vol. 49, No. 115, June 13, 1984, p. 24375.
19. S. B. Hudson. *Handbook of Application of Statistical Concepts to the Highway Construction Industry*, Part 1. National Technical Information Service, PB-220594, 1971.
20. Standard R9-86: Acceptance Sampling Plans for Highway Construction, Method A. In *AASHTO Standard Specifications for Transportation Materials and Methods of Sampling and Testing*, Part I. AASHTO, Washington, D.C., 1986.
21. W. J. Dixon and F. J. Massey. *Introduction to Statistical Analysis*, 3rd ed. McGraw-Hill, New York, 1969.
22. R. M. Weed. Bounds for Correlated Compound Probabilities. *Journal of Quality Technology*, Vol. 14, No. 4, Oct. 1982, pp. 196–200.
23. *Advancement of SQA and TQM in Design and Construction*. Study 7490. New Jersey Department of Transportation, Division of Research, Trenton, in preparation.
24. J. L. Burati, C. E. Antle, and J. H. Willenbrock. Development of a Bayesian Acceptance Approach for Bituminous Pavements. In *Transportation Research Record 924*, TRB, National Research Council, Washington, D.C., 1983, pp. 64–71.
25. J. Sweet. *Legal Aspects of Architecture, Engineering, and the Construction Process*, 2nd ed. West Publishing Company, Saint Paul, Minn. 1977.

PART 4: PLAN OF ACTION

The purpose of Part 4 is to build upon the material presented in Parts 1–3 to formulate a sound and workable plan of action that will significantly increase the effectiveness of transportation quality assurance practices nationwide.

LEADERSHIP NEEDED

If this goal is to be achieved, effective leadership is essential. The fundamental values of the leaders will be reflected in the performance of their subordinates. If the leadership is determined that the organization will produce an excellent product (whether it be roads and bridges or the standards and specifications that govern these items), then that product will tend to be of consistently high quality. Conversely, if the leaders neither appreciate nor understand the value of high quality, one can be sure that producing a quality product will not be foremost in the minds of their employees. If the full benefits of SQA are to be realized, leaders must take the following steps:

1. First become aware that present SQA practices are well behind the state of the art, exacting a very real price in terms of wasted effort and false security.
2. Then make the commitment to bring the necessary talent and technology to bear upon these issues.
3. Communicate this commitment to all members of their respective organizations and create an organizational culture that both fosters and demands excellence in what it does.
4. Arrange for competent training in basic SQA methods and encourage at least a few members of their organizations to acquire more advanced training.
5. Develop qualified staff and establish the necessary administrative procedures to ensure that their goals will be met.
6. Direct that all quality assurance applications are to use modern, state-of-the-art methods.
7. Reexamine the process by which voluntary consensus standards are developed to ensure that fundamental quality assurance principles will be applied correctly.
8. Take appropriate administrative action to ensure that the quality assurance program has credibility, both within the agency it serves and among the industrial organizations whose work it governs.
9. Provide visible and active support and set the direction for continued improvement.

If these steps seem to put the burden for resolving the problem squarely on the shoulders of transportation leaders, this is exactly what is intended. Quality assurance is not a problem that can be solved by lower or middle management, even with the best of intentions. If quality assurance is to be applied effectively in the transportation field, some distinct changes in both attitude and policy—backed up by the necessary resources—will have to come from the highest levels. In the remainder of this part of the paper we will elaborate on this obligation and how it might best be met.

1. BECOMING AWARE

It is almost axiomatic that before a problem can be solved, there has to be an awareness that the problem exists. It was noted in Part 1 that many current quality assurance standards and specifications are far from optimal, and Part 2 provided several examples of this. The use of inferior methods undermines the effectiveness of any quality assurance program and the long-term cost—based on the number of construction items involved nationwide—may be enormous.

Those who have begun to read the total quality management (TQM) literature know that this material is generally critical and disturbingly accurate. Collectively, these writers are sounding a warning that national leaders would do well to heed. If the United States is to avoid falling even farther behind in the ability to create and apply modern technology, Americans must be prepared to learn from those who have a clearly demonstrated superiority in this area. Those in the transportation profession seem to be open-minded enough in some cases—as evidenced by the recent European Asphalt Study Tour—and a similar approach must be taken with transportation quality assurance. The only difference is that the necessary statistical tools are already predominantly an American invention; all that is needed is to cultivate the good sense to use them to their fullest advantage.

But before this advice can be acted upon, it must be heard by leaders who are in a position to initiate the necessary changes. A means must be found to persuade top-level management to read this material and give it the thoughtful consideration it deserves. In our opinion, all transportation leaders who have an involvement with quality assurance should read Crosby (*1*), Deming (*2*), and Juran (*3*) as a minimum, with the primary emphasis being to gain an appreciation of

the cultural values and work ethic that have proven so successful for the nation's foreign competition. These leaders must then take a hard look at how the transportation profession measures up in this respect. If their assessment is made objectively, we believe that they will begin to see the need for major reforms.

2. MAKING THE COMMITMENT

Ever since a few pioneering states first began experimenting with SQA in the 1960s, transportation leaders at various conferences and conventions have eloquently expressed all the usual platitudes regarding the benefits of quality assurance. This must have been effective, because about three-fourths of the states now use SQA to varying degrees. However, when viewed from the standpoint of efficiency and effectiveness, the picture is very different. The transportation profession has hardly progressed beyond where it was 20 years ago, and current practices lag well behind the state of the art.

What is needed today in order for truly effective quality assurance practices to be established nationwide is a greater level of commitment. To be meaningful and effective, it must come from the highest levels—probably from within such organizations as AASHTO and FHWA—and it must involve more than reiterating the old platitudes. We believe that AASHTO, FHWA, and individual transportation agencies as well should voluntarily undertake a thorough examination of their own quality assurance activities with an earnest intent to seek out weaknesses and remedy them. This action, if conscientiously pursued, almost certainly would yield enormous long-term benefits in terms of better quality and better performance of the nation's highways and bridges.

3. CHANGING THE ORGANIZATIONAL CULTURE

One of the most difficult things to change is the way people think. But if sound quality assurance practices are to be established nationwide, this will require new thinking, new procedures, and a willingness to "work smarter" by discarding the older, less efficient methods with which many transportation agencies have grown comfortable. For some, these changes may seem to be quite radical.

Because this will be a significant departure from what most agencies are now doing, it must have the full cooperation and support of top management. First, the concept must be presented in a general way to explain why the changes are necessary and how they will benefit the organization and its employees. An increasingly obvious reason for such changes is the fact that the majority of states are experiencing serious financial shortfalls (*4*), making it necessary to use more effective methods and eliminate as much waste as possible. When construction quality is managed more effectively, this is clearly beneficial to both the transportation agency and the public it serves. Next, a series of training sessions on SQA must be conducted. Well-designed courses will explain the importance of the various techniques, present several examples of statistical acceptance procedures, demonstrate the superiority of correct methods, and provide numerous useful technical aids. Most employees will appreciate that this training provides them with additional marketable skills. And finally, management must remain actively involved to allay the concerns of the construction industry and deal with any opposition that might be encountered.

But management has an additional responsibility besides the duties just outlined. If an effective quality assurance program is to be established and maintained, it must have credibility both within the transportation agency it serves and among the various contractors' associations that it affects. To build a reputation for both competence and fairness, management must place qualified individuals in charge of these activities and empower them to apply their knowledge appropriately. In this way and in other, more subtle ways, management must create the type of professional environment that is conducive to the development of a first-rate program.

4. ESTABLISHING EFFECTIVE TRAINING PROGRAMS

Effective training accomplishes two very important objectives. In addition to providing the necessary skills to perform the assigned tasks, it provides the technical understanding necessary to instill a desire to do the job well. Employees who understand the importance of their function are more likely to be conscientious in performing it.

There have been some excellent training programs in transportation quality assurance (*5–7*) and a new one is in the developmental stages (*8*). Although these programs have been extremely useful in providing a familiarity with the basics of quality assurance, it has generally been beyond their scope to deal with the issues addressed in this paper—to encourage the use of the most effective, technically sound methods and to discourage the use of many current practices that are less effective and, in some cases, technically unsound (*9,10*).

The following quote from Deming indicates the importance he places on effective training in quality assurance:

> American management have resorted to mass assemblies for crash courses in statistical methods, employing hacks for teachers, being unable to discriminate between competence and ignorance. The result is that hundreds of people are learning what is wrong. . . . I make this statement on the basis of experience, seeing every day the devastating effects of incompetent teaching and faulty application. (*2*)

These are strong words from a highly respected source and they warrant careful consideration. Although there might be some measure of consolation in the fact that the private sector seems to be having similar difficulties in applying statistical principles correctly, this will not resolve the immediate problems with transportation quality assurance. Instead, consideration must be given to how current educational efforts can be supplemented and improved.

Deming makes an extremely valid point about the caliber of instruction. Although it might be thought that extensive statistical credentials are not required to teach an introductory course in SQA, this belief may be particularly shortsighted. Beginners, perhaps more than advanced students, need lucid explanations and careful guidance in the proper application of statistical principles. Like any mathematics course, the basics

must be clearly understood before the student can move on to more advanced concepts. In our opinion, the primary reason that current practices lag so far behind the state of the art is that qualified statisticians have not been routinely involved in either the development or the teaching of transportation quality assurance. The solution to this part of the problem is simple enough—leaders and managers must insist that future educational efforts in SQA involve all the necessary technical professions.

What should a comprehensive course in transportation SQA cover? As a minimum, it should include most of the fundamental concepts that are presented in Part 3 of this paper, with a special emphasis on the concept in Section 6 on operating characteristic (OC) curves. If it had been routine practice to construct OC curves, most of the problems cited in the section headed Inadequate Procedures and Practices in Part 2 could have been avoided.

5. DEVELOPING QUALIFIED STAFF

The effectiveness of a quality assurance program will be strongly dependent upon the knowledge and ability of those who administer it. Since top-level management cannot be expected to concern themselves with the day-to-day operation of the program, it is essential that they place competent staff in charge of these operations. This usually poses no problem as far as design, construction, inspection, and testing are concerned, because transportation agencies already have many employees well trained in these areas. A critical area of expertise that has been neglected, however, is statistical engineering. As already noted in the discussion of Step 4, we believe that this oversight is responsible for the many substandard quality assurance practices in existence today. Just as bridge design requires at least one member of an organization to have extensive training in structural analysis, the development of sound statistical specifications requires that at least one individual have a thorough understanding of applied statistics.

Because this was recognized as a vital prerequisite for a sound quality assurance program, the New Jersey Department of Transportation (NJDOT) created a new set of job specifications—the Statistical Engineer Series—to provide this necessary area of expertise. It was made an engineering series because the individuals in these positions must have a thorough understanding of the design and construction activities with which the SQA procedures must be coordinated. Beyond this basic knowledge, a substantial number of credit hours in statistical theory, acceptance sampling, and computer science is required.

Although the Department has not found it necessary to fill all the positions at any given time, the Statistical Engineer Series consists of three different positions, ranging from entry level to supervisor. This allows entry-level staff to advance as they gain more education and experience. Table 1 gives the educational and experience requirements for these positions.

The success of the NJDOT quality assurance program has been due in large measure to the in-house expertise provided by this group. The Department believes that the specifications developed (using the concepts outlined in Part 3) are practical, effective, fair to both parties, and legally defensible. In addition to developing statistical acceptance procedures, the statistical engineering group is responsible for preparing and conducting training courses on SQA methods, making presentations to contractors' associations and fielding the many technical questions that arise, giving expert testimony when required, and providing statistical assistance to many units throughout NJDOT. In our opinion, the decision to create a specialist group such as this is one of the strongest indicators of an agency's commitment to develop a sound quality assurance program.

TABLE 1 REQUIREMENTS FOR STATISTICAL ENGINEER POSITIONS

	Statistical Engineer Series		
	Entry Level	Intermediate Level	Supervisor
Engineering degree	BSCE or equivalent	BSCE or equivalent	BSCE or equivalent
Applied Statistics (credits)	12	18	24
Computer Science (credits)	6	6	6
Experience (years)	2	3	4

6. APPLYING THE TECHNOLOGY

It is well recognized that in order to achieve a permanent solution, it is necessary to identify and attack the root cause of a problem, not just the symptoms. For the problems associated with transportation quality assurance, the root cause clearly is not technical, because all the necessary technical tools are readily available. The three basic types of acceptance plans described in Section 13 of Part 3 utilize well-established, technically sound statistical procedures. They are suitable for virtually all transportation applications and could hardly be simpler to understand and apply.

To quote Crosby (*1*), "If something is easy to understand and makes sense, and yet isn't always done, there has to be a reason for it." He goes on to state that the usual reason is that management does not understand the value of a high-quality operation. Deming (*2*) laments, "Practically all of our major corporations were started by technical men—inventors, mechanics, engineers, and chemists—who had a sincere interest in quality of products. Now these companies are largely run by men interested in profit, not product."

This leads us to conclude that the real root causes are primarily cultural, behavioral, and managerial—precisely the topics that have received so much attention from the TQM writers—and this emphasizes all the more strongly the need for top-level management to begin to read and comprehend this material. If the benefits of sound quality assurance practices are to be realized, management has got to do more than sit idly by and condone current practices. To paraphrase some additional advice from Crosby (*1*), it is essential that management have some understanding of the quality assurance process in order to have a realistic sense of what can and should be done.

7. IMPROVING TECHNICAL STANDARDS

In our opinion, much of what has been published in the way of national standards and guidelines on SQA in the transportation field is woefully inadequate. Some of the procedures have no rational mathematical basis (see Example 4 under Inadequate Procedures and Practices in Part 2) and most have not been subjected to any formal test to demonstrate their adequacy (see Section 6 on OC curves in Part 3).

Both Crosby and Deming make comments that, although directed at quality management in general, can be seen to apply to the standards-writing process in particular. First Crosby:

> Top managers may or may not realize what has to be done to achieve quality. Or worse, they may feel, mistakenly, that they do understand what has to be done. Those types can cause the most harm. (*1*)

Then Deming:

> Best efforts are essential. Unfortunately, best efforts, people charging this way and that way without guidance of principles, can do a lot of damage. Think of the chaos that would come if everyone did his best, not knowing what to do. (*2*)

If national quality assurance standards are to be improved, the process by which they are developed must be improved. It would be highly desirable, of course, for committee chairmen to be sufficiently knowledgeable in SQA procedures to know the level of competence that should be insisted upon. Lacking this, they must at least be leaders who will demand technical competence in anything for which they are responsible. This will require seeking out the necessary expertise, even if it means going outside the committee membership. It will also require inviting an open and thorough scrutiny of the finished product. This can best be accomplished by arranging for an independent review by qualified individuals to ensure that standards are technically sound before they are balloted. These are not difficult steps to accomplish and it is up to high-level leaders to insist that this approach, or something very similar, be made an integral part of the standards-writing process.

Both the correct development and the review of any SQA standard can be further facilitated by requiring the underlying mathematical principles to be explicitly covered in an appendix. This forces the writers of the standard to fully grasp the subject matter and it enables the reviewers to know precisely what the writers had in mind. An excellent example of this is in Method A of AASHTO Standard R9 (*11*).

A recent example illustrates how successful this overall approach can be. A new sequential sampling standard was prepared by one of us as part of a technical support group for AASHTO Technical Section 5c, Quality Assurance, Data Evaluation, and Acceptance Plans. The draft was then sent to two reviewers, also noncommittee members, who had expertise in this area. Although no major errors were found in this particular case, a number of minor errors were corrected that probably would have escaped detection had the balloting process been relied upon to provide the technical review. As a final safeguard, the draft was also sent to TRB Task Force A3T51, Statistical Methods in Transportation, for further review. A comprehensive and thorough approach such as this, which involved both engineering and statistical expertise in both the writing and review stages, has almost certainly produced a valid and effective standard that can be used with confidence. We remain convinced that had a process such as this been used routinely in the past, most of the problems in existing quality assurance standards would have been avoided altogether.

8. ESTABLISHING CREDIBILITY

As already noted in the discussion of Step 3, the long-term success of a quality assurance program will be strongly dependent upon its degree of credibility, both within the transportation agency and among the industrial organizations whose work it governs. People will generally support (or at least offer less resistance to) programs that seem to make sense and in which they have confidence. Three essential factors in achieving credibility are effectiveness, administrative thoroughness, and fairness.

To be perceived as effective, quality assurance programs must be applied to construction items that are clearly of importance and for which the failure to achieve high quality is known to be costly, either in terms of economics or safety. They need not necessarily be items that have an obvious quality problem because, as discussed in Part 2, serious quality problems often do not become evident until several years after construction. A second requirement for effectiveness is a demonstrated improvement in those properties that are believed to be closely linked to performance. For example, if quality assurance programs tend to encourage stronger concrete, thicker pavements, or greater uniformity in a variety of construction processes, they will be perceived as effective. As also noted in Part 2, results of this sort have typically been observed when quality assurance programs have been implemented.

By administrative thoroughness we are referring to the degree to which the transportation agency is attentive to the various details that are characteristic of a carefully conceived implementation plan. These details include the following:

1. Presentations by top management to outline organizational goals;
2. Broad training to acquaint design, construction, and materials personnel with SQA methods;
3. Specialized training for a few individuals (such as those in the statistical engineer positions);
4. Meetings with contractors' associations to give them advance notice of the new procedures that are being developed;
5. The thorough testing (usually by computer simulation) of all prototype statistical acceptance procedures before actual field trials are attempted;
6. The development of instructional bulletins and other technical aids to guide field personnel in day-to-day operations;
7. The scheduling of small pilot projects, possibly with relaxed pay-adjustment plans, to field-test the new specifications;
8. Feedback sessions, conducted separately with agency and contractors' personnel, to evaluate the outcome of the field trials;

9. The prompt correction of any deficiencies and, if major changes are involved, the rescheduling of additional field trials; and

10. Broad implementation once the pilot projects have been demonstrated to be successful.

NJDOT has used this approach for several years and found it to be effective. At the very least, it will minimize the difficulties normally associated with the implementation of a new SQA specification. It increases the likelihood that the initial procedure will be technically sound and effective, it allows time for all parties to become familiar with the new program, and it provides a mechanism to address any concerns that transportation agency and contractors' personnel might have.

The third factor in establishing credibility is fairness. SQA specifications must protect the interests of the transportation agency and at the same time treat the contractor in a fair-minded fashion. Only through the development of OC curves (explained in Section 6 of Part 3) is it possible to determine whether the acceptance procedure will properly accept good quality and reject poor quality. It also is necessary to communicate to the contractor what level of quality is desired, and when that level of quality is consistently supplied, the contractor has a right to expect an average pay factor of 100 percent. In a recent paper (*12*), it is noted that many existing specifications fail to accomplish this and an explanation is given for how bonus provisions provide a convenient way to correct this flaw. Still another aspect of fairness relates to the amount of payment withheld when the quality is substandard. A rationale believed to be both equitable and defensible is described in Section 10 of Part 3.

Although the attitude of the majority of contractors toward SQA is still very apprehensive, there definitely has been a mellowing in recent years (*13*). If additional progress can be made toward the establishment of a consistent national policy—using only procedures that have been proven to be effective, efficient, and fair—it is possible that the construction industry might eventually see this as beneficial to them in that it would provide a consistent basis for both bidding and production. At the very least, it would enhance the credibility of transportation agencies and SQA practices in general.

9. PROVIDING FUTURE DIRECTION

One of the "managerial myths" cited by Juran (*3*) is that quality will get top priority if upper management so decrees. He goes on to state that nothing of the sort will happen unless management follows through with fundamental changes: specific goals, the resources necessary to meet those goals, realistic timetables, and an administrative process to make sure the goals are met.

One way the problems in transportation quality assurance could be resolved would involve a fundamental change in the Intermodal Surface Transportation Efficiency Act (ISTEA). In its present form, it provides for the management of pavements, bridges, congestion, and safety. Conspicuous by its absence is any provision for the management of quality. To quote a statement from a recent FHWA Administrative Memorandum (*14*), "Very simply, research can only be effective if it finds its way to the road." One of the best ways to make sure that present and future research results do reach the road is to establish thorough quality assurance programs that will guarantee that what is called for in the plans and specifications is, in fact, received. The most effective way to bring this about would be to make quality the fifth management mandate in a future version of the ISTEA.

One of the directives in the newly enacted ISTEA calls for the Secretary of Transportation to undertake a study of state procurement practices, including statistical acceptance procedures, and to prepare a report that, among other things, will examine the need for a national policy on transportation quality assurance. This is an important step in the right direction, but if it is to be effective, we think it should be strengthened by specifically requiring the study group to be composed of individuals from the transportation, academic, and industrial fields who have specialized training in SQA. Both Crosby (*1*) and Deming (*2*) caution in the early pages of their books that the results of such an effort may be less than useless if it is not conducted by individuals having the necessary expertise.

As noted in the discussion of Step 7, the way to improve quality assurance standards is to improve the process by which they are developed. This is generally true of all aspects of quality assurance. The problems in the transportation profession today are largely the result of faulty processes, and to correct these problems, the processes themselves must be changed. If new efforts—whether they be standards, specifications, or training programs—are to achieve their full potential, they must be more than carbon copies of the past. If leadership has been ineffective or nonexistent, then new leadership must be found that can provide the direction necessary to reach these goals.

What can individual leaders do? For one thing, they can decide to become more knowledgeable by reading current literature on TQM and SQA. They can then begin to discuss these issues with their counterparts in other states and organizations, building the vocal advocacy necessary to get this topic on the agendas of various policy-making and law-making bodies. Ultimately, if a sufficient number of transportation leaders can be convinced to focus their attention on the issue of transportation quality assurance, we believe that they will realize the need for a consistent national policy.

SUMMARY AND CONCLUSIONS

In Part 1 of this paper we noted that the highway system represents an investment valued in trillions of dollars and that any measure that could improve its performance by even a few percent would result in savings of billions of dollars. SQA—currently in use or under development in approximately three-fourths of the states—has proven to be a very effective tool to encourage high-quality construction. Unfortunately, there is great disparity in the manner in which it is applied from state to state, and many current practices are far from optimal. It was concluded that sweeping reforms are needed and that a consistent, scientifically sound national policy on transportation quality assurance should be established.

In Part 2 we warned of several obstacles—technical, managerial, political, and cultural—that could impede either the effective application of SQA procedures by individual agen-

cies or the establishment of a sound national policy. Perhaps the biggest obstacles are a general lack of awareness of just how far behind the state of the art transportation quality assurance really is, the failure to insist that those involved with quality assurance be thoroughly educated in these matters, and a work ethic that seems to be devoid of any real pride in what it produces.

In Part 3 we presented a series of fundamental concepts in the belief that reliable quality assurance technology must be based on sound scientific, mathematical, and legal principles. These principles, proven by actual field application over a period of approximately 20 years, provide the basic technical building blocks from which practical and effective quality assurance programs can be developed. The guidance they provide will be useful both to individual agencies and to a task force contemplating a national policy on transportation quality assurance.

In Part 4 we outlined a series of steps that transportation leaders must take if truly effective quality assurance practices are to be established nationwide. They must first read the appropriate literature to begin to understand how backward many current practices are and to learn that better practices can produce significant savings in both cost and performance. They must then make a firm commitment to bring about the necessary changes and begin to create an organizational culture that understands and appreciates the value of total quality management (TQM). They must arrange for the necessary education and training, which is an absolutely essential part of such a program. And finally, they must set a reasonable timetable for the conversion to state-of-the-art methods, establish the necessary administrative procedures to keep this effort on track, and remain actively involved to lend support when it is needed.

At the national level, responsible leaders have an obligation to thoroughly explore the benefits that might be derived from a more rigorous application of modern quality assurance methods. They must insist that this topic be placed on the agendas of policy-making and law-making bodies and be made the subject of serious study by individuals with the appropriate technical expertise. We believe that such a study will conclude that a national policy on transportation quality assurance will pay huge dividends in terms of better quality and better performance of the nation's highways and bridges.

REFERENCES (Part 4)

1. P. B. Crosby. *Quality Is Free*. McGraw-Hill, New York, 1979.
2. W. E. Deming. *Out of the Crisis*. Massachusetts Institute of Technology, Cambridge, 1986.
3. J. M. Juran. *Juran on Leadership for Quality*. The Free Press, New York, 1989.
4. R. Thomas and E. Thomas. Where Did All the Money Go? *Newsweek*, July 1, 1991, pp. 25–28.
5. J. H. Willenbrock. *Statistical Quality Control of Highway Construction*, Vols. I and II. FHWA, U.S. Department of Transportation, 1976.
6. *Demonstration Project 42: Highway Quality Assurance*. Demonstration Projects Division, FHWA, U.S. Department of Transportation, 1976.
7. J. L. Burati and C. S. Hughes. *Highway Materials Engineering, Module I: Materials Control and Acceptance—Quality Assurance*. National Highway Institute, FHWA, U.S. Department of Transportation, 1990.
8. *Demonstration Project 89: Quality Management*. FHWA, U.S. Department of Transportation, under development.
9. R. M. Weed. Revision of a Flawed Acceptance Standard. In *Transportation Research Record 1056*, TRB, National Research Council, Washington, D.C., 1986, pp. 21–35.
10. P. A. Kopac. Current Practices in Acceptance of Bituminous Concrete Compaction. In *Transportation Research Record 986*, TRB, National Research Council, Washington, D.C., 1984, pp. 43–46.
11. Standard R9-86: Acceptance Sampling Plans for Highway Construction, Method A. In *AASHTO Standard Specifications for Transportation Materials and Methods of Sampling and Testing*, Part I. AASHTO, Washington, D.C., 1986.
12. R. M. Weed. Adjusted Pay Schedules: New Concepts and Provisions. In *Transportation Research Record 986*, TRB, National Research Council, Washington, D.C., 1984, pp. 32–38.
13. M. L. Powell. Quality—Contractors Cannot Afford Less. *TR News*, Nov.-Dec. 1986, pp. 38–43.
14. *Marketing Highway Technology and Programs*. Administrative Memorandum. Office of Research and Development, Operations, and Support, FHWA, U.S. Department of Transportation, Oct. 1990.

Can Total Quality Management Work in the Public Construction Arena?

THOMAS H. OSWALD AND JAMES L. BURATI, JR.

Total quality management (TQM) has been demonstrated to increase productivity and profitability in numerous applications in manufacturing and some service industries. However, its viability as a major management philosophy in engineering and construction has not been widely established. The same is true in the general area of government agency operations and more specifically in the field of public-sector construction. In this paper we discuss the principles and operational elements of TQM in terms of the unique characteristics of public construction. The quality improvement programs of three major federal construction agencies are discussed, and conclusions are drawn concerning the success factors for making TQM work in public construction.

Total quality management (TQM) is the name given to a broad approach to the improvement of the quality and productivity of manufacturing, service, educational, and many other forms of institutional activity. TQM is based on "new" concepts of

- Customer-supplier relationships,
- Employee involvement in decision making,
- Teamwork,
- Rigorous analysis of work as a process,
- Statistical quality control, and
- Transformation of management from a focus on controlling to a focus on leading.

The "quality" being subjected to these techniques defies concise definition. It includes

- Satisfied customers;
- Repeat business, including long-term business relationships;
- Reduction of variation and rework in work processes;
- Reduction of production-service cycle times;
- The absence of disputes; and
- Greater alignment of individual and corporate objectives through enlightened manager-leadership.

TQM has made enormous strides in the United States since it became known and was in instant demand in the late 1970s and early 1980s. A wave of interest in, and serious implementation of, these "modern" philosophies and techniques of management has swept across the country, accomplishing in less than 15 years what has been evolving in Japan for nearly 40 years.

T. H. Oswald, 201 Lakeside Court, Clemson, S.C. 29631. J. L. Burati, Jr. Civil Engineering Department, Lowry Hall, Clemson University, Clemson, S.C. 29634-0911.

This is not an indication that this country was more adept at recognizing and developing TQM than was Japan. Rather, the fundamentals of the new "art-science" of management were worked out during a long period of time in Japan, whereas U.S. industry at the same time, but erroneously, perceived that all was well and quality was "good enough." Growing foreign competition coupled with internal awareness of declining productivity eventually got the attention in the United States. Some fortuitous exposure to the reasons behind Japan's advances in quality caused a surge of interest in bringing these methods to this country (*1*).

Since 1980 there has been extensive emulation of Japan's total quality practices by U.S. industry. An early but unsuccessful foray by many American companies was in quality control (QC) circles, which generally failed in the United States because of the failure of management to implement them within a top-to-bottom, carefully structured and integrated quality effort embracing not only people, but cultures and philosophies, policies and processes, teamwork and training. Similar experiences resulted from ill-conceived attempts to "install" statistical quality control (SQC) as the necessary and sufficient methodology to solve all quality problems.

As the trans-Pacific learning experience progressed, American managers, primarily in the manufacturing industries, came to understand the meaning of the "total" in total quality management, that is, its impact on management as well as labor, techniques as well as tools, methods as well as materials. By the end of the 1980s, Americans were beginning to see and to satisfy the necessity for an all-out approach in the implementation of TQM.

The total quality movement eventually spread from the manufacturing sector to the service industries where, as with manufacturing, U.S. managers not only successfully implemented the "Japanese version," but also made the necessary cultural adjustments and began to assemble a distinctly American TQM body of knowledge.

The U.S. implementation of TQM is now well into the institutionalization phase. Quality has itself become a growth industry, as evidenced by the proliferation of books, articles, seminars, consulting services, conferences, and congresses under some banner or other involving "quality." In addition to these price- or fee-based manifestations, there has developed a large body of scholarly research and development both by academia and by industry. The federal government has instituted and awarded the Malcolm Baldrige National Quality Award and numerous other awards centered in companies, associations, and *even* in the federal government. In short, TQM is very much alive and in relatively good health in the United States.

OBJECTIVE

In this paper an attempt is made to address the nonessentiality of the word *even* in the next-to-last sentence of the preceding section. Why now, 10 years into the "U.S. quality experience," is it necessary to single out the federal government, any level of government for that matter, as somehow different in regard to the applicability of TQM? The same may be asked about the U.S. construction industry. Is TQM really for manufacturing and mass-market service industries only? Combining *government* and *construction*, we get *government* (or public-sector) *construction*, the topic covered in this paper. Is it reasonable to believe that TQM can work even in public construction? Many skeptics do not think so.

If we have been successful, this paper answers that question. We look at some of the foundation principles and elements of TQM and point out the pros and cons of each regarding government or construction or public construction. We cite the activities of three public agencies whose missions are heavily construction oriented, and that have put in place innovative and productive improvement systems that directly or indirectly are based on TQM principles.

CASE FOR TQM

At this stage in the growth of TQM in the United States, it is unnecessary to go to great lengths to cite the reasons why this system of management should be considered by any business or similar organization wishing to maximize its mission effectiveness. Trade literature, the literature of quality-based associations such as the American Society for Quality Control (ASQC), and general business and professional literature sources contain increasing numbers of articles on the proven benefits of adopting a comprehensive and determined management approach to quality. The so-called "Deming chain reaction" (*2*) starts with "improving quality" and proceeds through decreasing costs and increasing productivity to increasing market share and staying in business, and finally results in "providing jobs and more jobs." Variations of this sequence have been validated over and over in several major sectors of U.S. industry. The publicity surrounding the Malcolm Baldrige National Quality Award provides widespread testimony to the fact that TQM works.

But does it work in construction? Or in government or government (i.e., public-sector) construction?

Government Progress

There is an increasing level of interest and activity among public agencies, mainly at the federal level, but with several notable examples at the state and municipal levels as well. The Federal Quality Institute (FQI) was chartered in 1988 to provide a source of TQM information and to assist in catalyzing action by federal agencies in implementing TQM. FQI has assisted numerous federal organizations, and has documented and publicly recognized exemplary achievements by such agencies as the Cincinnati Service Center of the Internal Revenue Service, the Defense Industrial Supply Center, the Naval Air Systems Command, and the National Aeronautics and Space Administration Lyndon B. Johnson Space Center (*3*).

State and municipal agencies also are active in the pursuit of TQM. Holt (*4*) and Hughes (*5*) discuss quality management in the Minnesota Department of Transportation (Mn/DOT). The Madison (Wisconsin) Area Quality Improvement Network (MAQIN) is a good example of what can be done at the local level by visionary and determined leaders (*6*). Walton describes two highly successful agency- and areawide TQM programs in her books on the management methods taught by W. Edwards Deming: The Philadelphia Area Council for Excellence (PACE) (*7*) and the Departments of Navy and Defense (*8*).

Construction Industry Progress

Research performed in 1988–1989 by the Construction Industry Institute (CII) (*9,10*) showed a high level of TQM awareness and initial organizational preparation by a number of leading U.S. engineering and construction companies and by major purchasers of their services. Our follow-on CII research, to be published in late 1992, indicates that the engineering and construction industry is continuing to make great strides in implementing TQM. The industry is not merely copying the techniques of manufacturing, but is innovatively adapting the new art-science to the unique characteristics of the project-based, multiparty construction process. Engineering and construction, especially in the upper range of company and project size, is indeed taking its rightful place in the implementation of TQM in the United States. Although no research is available to show a corresponding level of activity in smaller companies and project sizes, numerous individual efforts are known to exist. It is our opinion that small-project TQM will spread rather rapidly during the next few years as knowledge of its benefits is disseminated and affordable implementation models evolve.

PRINCIPLES

The fundamental principles on which TQM is based are customer satisfaction and continuous improvement. The examples cited in this paper are motivated by a desire to bring about the fullest degree of satisfaction of the government agencies for whom the Federal Highway Administration, the Corps of Engineers, and the Naval Facilities Engineering Command are procuring design and construction services. To the extent that their customer agencies have developed projects best meeting the needs of the ultimate customer-user—that is, the public—TQM efforts will pass through the supplier-customer chain with the greatest benefit to the taxpaying public.

Continuous improvement on a macro scale is fostered in the three agencies by (a) the careful deployment of TQM awareness and emphasis throughout the organizations, extending to their relationships with their contractors; (b) the internal commitment of resources to TQM; and (c) institutionalization of TQM by the agencies' membership and active participation in quality associations and other public activities.

ELEMENTS OF TQM AS APPLICABLE TO PUBLIC CONSTRUCTION

Some of the generally accepted elements of TQM are here discussed in terms of our perceptions of how public construction should (and perhaps does) effectively implement them.

Management Commitment and Leadership

TQM is not a system of mechanical processes with a finite beginning and end that can be "installed," executed, and retired, the benefits having been permanently captured for the ongoing good of the organization. It is, rather, a culture and philosophy that must permeate the organization as The method of management (*9*). The leaders of the agencies included in this paper obviously have this commitment and understanding of true leadership. Other agencies as well as many private companies have tried and failed, or have not been allowed to seriously consider implementing TQM primarily because their senior managers were ignorant of, or were actively opposed to, the innovation that it entailed (*11*). Conventional objections include: "We already do quality work," "This will be too costly," and "It will never work here; this is not Japan." In many instances, field activities of federal agencies have succeeded in TQM on their own local initiative, sometimes with no strong support from Washington. The key success factor, after the education of senior management, is the degree of individual innovation and initiative inherent in their personal style of managing.

As long as public agencies (properly, in our opinion) encourage and reward voluntary contractor TQM, those contractors who actually receive work are most likely to be those whose own management is enlightened and committed to TQM, or who are at least willing to try it. If procurement agencies begin to codify TQM requirements in a "checklist" sense, there will probably be an erosion of the long-term benefits of TQM to the using public. Some of the early Contractor Quality Control (self-inspection and certification) programs allowed poor contractors to meet the letter of the requirements while sowing the seeds of later facility failures by not having CQC as an integral part of the way they did business. For TQM to succeed in public construction, therefore, top and intermediate management of both the public agency and the contractors, subcontractors, and suppliers must have the necessary level of TQM commitment and leadership.

The development of such commitment on the part of the contractors will be a function of the general business education and experience of the principals as they respond to real and perceived business success factors, including what the rest of the market is doing. On the part of the public agency the same is true, but another key influence enters the equation. This is the tendency of public agencies to experience "revolving door" leadership. When the senior managers are in the military or when heads of civilian agencies are politically appointed, serve briefly, or both, and go or return to industry, the "constancy of purpose" espoused by Deming (*2*) will be a special challenge. In all of these cases, only time and the continued "American institutionalization" of TQM will result in such leadership's being the rewarded norm, "commander after commander, secretary after secretary, and director after director."

In the meantime, the consistent practice and growth of TQM as The way to manage can be aided significantly by the knowledge and earnest support of the career civil service managers, who provide local continuity over long periods of time. Again, this will be a function of both the individual manager (initiative, innovativeness, sense of true leadership and of public stewardship, etc.) and how his or her efforts are rewarded by the system.

Process Emphasis

Most, if not all, public agencies can be characterized as a "network of processes." Their existence and missions are the result of legislative or executive (sometimes even judicial) actions, which are implemented by the writing of voluminous directives, policies, and procedures. This documentation of the way things are intended to work provides these organizations with a head start in prioritizing and analyzing their work processes with a view toward their improvement. The federal programs cited below are using the tools and techniques of TQM to improve the broad processes of procuring the design and construction of public projects. The criticality of process is doubly important in the principal work of these agencies, because it is performed through contracts, where not only the initial efficiency is enhanced by optimum execution of the processes, but the life-cycle effectiveness of the work is increased by the reduction of failures, disputes, and litigation.

Emphasis on process, therefore, is a fertile area for TQM implementation in public construction. It can produce significant, measurable benefit when applied internally by the parties and across their contractual agreements.

Training

Government agencies, especially those associated with the military, and construction industry organizations historically have understood and exploited the value of functional training. This also gives them a practical head start into the successful implementation of TQM. However, in addition to functional training, there must be a commitment to education in the philosophy and techniques of TQM. Public agency representatives are increasingly visible at seminars and "colleges" offered by professional quality consultants. In-house training in quality awareness and practices is also common in the more progressive agencies. To back up their programs of technical training and continuing education with the quality knowledge needed to achieve long-term improvement of their organizational performance, design and construction firms likewise are availing themselves of these educational resources.

The main challenge for public and private organizations alike is to ensure that the formulation, delivery, and subsequent use of quality training take place only with the objective of bringing about pervasive and permanent modification of the way business is conducted. "Training" in only the super-

ficial aspects of quality improvement will have short-lived and disappointing effects.

Teamwork

Contractors as well as procurement agencies operate internally around their networks of processes, as discussed above. However, the parties in both sectors historically have organized according to function rather than process, with the same practices of "pass down the line" or "over the wall" conduct of business. Further, the traditional acceptance of the adversarial approach to contracting has produced many of the tired old paradigms of the impossibility of cooperative project execution in government work. The agencies discussed below are taking a serious approach to the improvement of ongoing processes, as well as to the solution of repeated problems, by organizing, training, and chartering cross-functional teams who are intimately involved with the processes under study. Likewise on the contractor side, the most successful TQM companies have rigorous, yet simple team approaches to achieving process improvement and problem solving.

The ultimate form of teamwork in construction is that which occurs between traditionally adversarial parties looking at each other through the filtering effect of a contract. Can such teamwork succeed in public construction? Clear, positive answers are present in the partnering examples of the FHWA and the Corps of Engineers.

Fact-Based Decision Making

The most common interpretation of this element is that in which the facts are statistical in nature. Such facts are the most incontrovertible for the purpose of decision making, provided they have been developed using statistically valid methods of data collection and analysis. The statistical methods used in materials acceptance testing and other applications, notably in highway construction and for structural materials, have long been standard management tools in construction. Much formal and informal research is currently being performed by owners, designers, and constructors to formulate measures of performance in other areas, which are both statistically valid and meaningful for continuous improvement. Most of these organizations are, properly in our opinion, exercising caution in attaching life-or-death importance to the expanded use of statistical methods. Appropriate statistical techniques do exist, but their adoption must be carefully approached, and the timing of their use should come at some distance into the TQM development of the organization.

Not all facts must be statistical. Efficient project execution depends on thousands of decisions, usually impinging differently on the interests of the respective contracting parties, and based on facts that may be seen just as differently by each party: Does the weather forecast justify cancellation of tomorrow's concrete delivery? Should we invoke the highest level of expediting for the cooling tower? Fact-based decisions on questions such as these require determinations of fact in advance. Because this is outside the psychic capability of most project managers, the decisions must be made based on the best professional judgment of the decision maker. Even so, unilateral decisions based on excellent professional judgment are often at the center of claims and disputes.

Enter TQM and teamwork or partnering. Given a formal predisposition to cooperate in the day-to-day execution of the project, a consensus process can produce a "fact" on which to base a joint decision. This fact and the resulting decision can be equally as wrong as a decision erroneously arrived at by a single party. Instead of a postmortem statement, "You made a bad decision; now what are you going to do about it?" we would more likely hear, "We guessed wrong this time, but we can work around this undesirable outcome." Consensus-based determination of "fact" and attendant consensus decision making based on the jointly arrived-at information have preserved the cooperative working relationship, and most important, they have precluded a potentially serious interruption in the momentum of the project execution.

Supplier Involvement

The main suppliers in a project are the various contractors and vendors. Partnering efforts with successful bidders, as well as prequalification of bidders, and legal, preferential treatment of design firms that have superior quality performance records, are examples of successful contractor involvement in one or more of the three federal agencies cited below. Prime contractors and subcontractors extend these principles to their selection of vendors and suppliers, with increasingly common emphasis on reducing variation through reducing the supplier base. Although public procurement regulations historically have seemed to discourage, if not to outright prohibit, cooperative effort between procurer and supplier, and also any significant reduction of the number of suppliers, nevertheless the regulations often present wide latitude for contracting officers to legally exercise innovation and initiative, observing both the letter and the intent of the regulations and simultaneously serving the public interest through TQM-based procurement approaches.

Customer Service

Customer satisfaction appears both as one of the fundamental principles of TQM (see preceding section) and here as the last of the elements of TQM practice. Dedication to customer service, both in principle and in practice, requires interjecting one's own organization into the stream of operational factors that determine one's customer's success. An organization or an individual usually has multiple customers, some internal to the organization, and some external. In all such relationships, the accurate communication of the customer's requirements (as opposed to just the passing of information) is the necessary prerequisite to all other TQM effort.

In public construction, as in private-sector industrial construction, the multiplicity of stakeholder organizations on the customer side of the project raises a special challenge for the procurement organization. There must be in place a consistent and reliable process for identifying the respective customers along the chain; defining their interests, needs, and expectations; and implementing a project management system that

preserves the focus on the several degrees of customer interests throughout the entire project. The "life cycle project manager" concepts of the Corps of Engineers and the FWHA are designed with this purpose in mind.

The contractor looks at the contracting officer of the public agency as its customer and, assuming that the contractor is enlightened in the practice of TQM, it should be able to satisfy that customer through at least a quality-based approach to contract work, but ideally through voluntarily availing itself of any offer by the government to enter into a partnering arrangement.

MAJOR IMPLEMENTATION EXAMPLES IN PUBLIC CONSTRUCTION

The potential for improvement in the conduct of the public's business in the acquisition of design and construction services has been recognized and exploited by several federal agencies. Three examples are presented here.

The point to be drawn from these examples is that the old paradigms of "government bureaucratic rigidity" and "inherent adversariness and mutual distrust" in government construction are now beginning to pass into history.

Federal Highway Administration (FHWA)

FHWA experienced heavy pressure in the 1980s from the Office of Management and Budget (OMB) to increase productivity. The Administration obtained the concurrence of OMB to pursue greater productivity by improving the total quality aspects of how it carried out its mission. With the support of its executive leadership, FHWA currently has an active TQM process in place, involving most of the major elements commonly attributed to successful TQM implementation, as discussed in the following sections.

Emphasis on Process Improvement

Several disciplines and activities are involved in the analysis and systematic improvement of key internal processes and those that affect customers.

Training

Training is concentrated, via a team focus, on quality awareness, identification of internal and external customer-supplier relationships, and quality improvement tools and techniques. The agency initially employed outside training consultants in the areas of team building and facilitation but has evolved to mainly in-house training, tailoring its training programs to best fit its needs.

Technology

Technology is used effectively to reduce or eliminate sources of error and variation, the two prime targets in quality improvement. Computer-assisted drawing and drafting (CADD), interactive automated field surveys, and advanced photogrammetry techniques are employed, as well as the latest technology in bridge and roadway design and construction, geotechnical engineering, and hydrology.

Fact-Based Decision Making

FHWA continues the tradition of using SQC methods in highway construction, including new techniques for evaluating pavement smoothness. Nonstatistical fact-based decision making enters the FHWA process in its practice of teamwork for process improvement and for conflict resolution.

Supplier Involvement

The practice of partnering on a specific project is available to FHWA contractors on a voluntary basis. Partnering in the context of public contracts does not mean the existence of "evergreen" multiyear, multiproject contracts, as is the case in many areas of the private sector, principally in industrial construction. Rather, FHWA offers, through the project solicitation, to enhance interparty relationships by formally constituting a high-performance joint project management team dedicated to the pursuit of mutual and respective objectives and the promotion of nonadversarial project execution. If the contractor elects to participate in the partnering process, a formal agreement is entered into, which lays out the respective parties' interests and provides for enhanced joint project management processes. Either side can terminate the agreement at any time for any reason. The parties meet shortly before the commencement of project work and, with the help of an independent facilitator, work out the partnering agreement. Typically the meeting includes the contracting officer, project engineers, supervisors, the contractor's key supervisors, a representative of its home office management, and sometimes subcontractor representatives. Attention in these meetings and in the subsequent partnering agreement is focused on the development of trust among individuals and parties, on communications, and on the promotion of nonadversarial relationships.

Contract Procurement Improvements

In addition to the partnering process, FHWA is implementing quality-supportive contractor selection methods. To increase the probability of satisfactory contractor performance, a two-step bidding process is used, which achieves these quality objectives while conforming to the requirements of the Federal Acquisition Regulations. This technique provides greater assurance of quality performance and therefore better satisfaction of the customer's—that is, the public's—requirements while adhering to regulatory requirements for award of public contracts to the bidder submitting the lowest conforming bid.

Customer Service

FHWA considers that its customer interests are best served by the provision of consistent project management over the

entire time span of the project. It assigns an experienced project manager and associated team members to manage the project from start to finish. This practice starts in a manner similar to the contractor partnering process, in which project scope and criteria are carefully worked out with the client agency, followed by consistent review and reinforcement of the user-customers' needs throughout the project.

Self-Assessment

FHWA evaluates its quality improvement process against the federal Quality Improvement Prototype (QIP) Award criteria, which are administered by the Federal Quality Institute. By so doing it ensures that it is focusing on the "total" aspect of total quality improvement and is applying appropriate measurement methodology to evaluate its success.

U.S. Army Corps of Engineers (USACE)

USACE in 1982 convened a select group of senior Corps officials to study and make recommendations concerning improvement of the quality of the Corps' construction, contractors' quality control, and the Corps' administration of the quality assurance function. This Blue Ribbon Panel on Management of Construction Quality in the U.S. Army Corps of Engineers submitted a number of recommendations and specific actions in its 1983 report. The Corps then developed and implemented a plan and has completed action on most of the recommendations contained in the report. The Blue Ribbon Panel's output provided the basis for major initiatives and directions in the Corps' pursuit of quality in engineering and construction. Although not identified specifically as a TQM effort, nevertheless the panel's recommendations centered on such items, identified in a 1987 implementation report (*12*), as

- Customer/user satisfaction,
- Lessons-learned system on design deficiencies and repetitive errors,
- Primacy of quality assurance and demand for performance,
- Rewarding good contractor performance,
- Improvement of interaction between engineering and construction elements, and
- Improving technology transfer from R&D to field offices.

As can be seen, these items (and many others recommended by the panel) fall under the TQM umbrella: customer satisfaction, process improvement, supplier (contractor) involvement, organizational relationships and communications improvement, and the use of technology.

More recently, specific initiatives to improve the Corps' military construction acquisition process include (Headquarters, USACE, unpublished listing):

- Establishment of requirements for design quality control plans by architect-engineer (AE) contractors and design quality assurance plans by each Corps district.
- Automated Review Management System (ARMS), a computer-based system to improve the preparation, submission, and resolution of design review comments, including feedback to reviewers.
- AE Responsibility Management Program (AERP), a formal management system that deals with Corps responsibilities and procedures in cases of AE failure to perform.
- Use of computer disk read-only memory (CD-ROM) to improve the dissemination, currency, data access time, and quality of Corps Guide Specifications, industry standards, and other criteria.
- Life Cycle Project Management (LCPM), a major initiative currently being implemented, which provides for single-point responsibility for management of the design and construction process of a project from inception through completion and turnover to the customer-user. LCPM goals are to improve project execution and customer involvement; increase accountability in the areas of scope, quality, budget, and schedule; and provide project management continuity.
- Various information and feedback systems designed to improve customer service, such as the Engineering Improvement Recommendation System (to capture field and customer improvement ideas), Design and Construction Appraisal System (to capture the results of inspections by Design and Construction Evaluation Teams), a Management Lessons Learned System, and a Construction Bulletin process that provides timely guidance on policy issues in construction contract administration.
- Partnering. The Corps, with the cooperation of the Associated General Contractors (AGC), is institutionalizing across all its construction operations the concept of partnering for improved joint project execution. As with the FHWA partnering implementation, described above, this concept constitutes a carefully conceived and deliberate, "new" approach to doing business with contractors. It is based on the creation of a formally structured, noncontractual agreement between the contractor and the cognizant Corps construction office. The partnership arrangement seeks to build trust among the major stakeholders in the success of the project and to work from a basis of cooperative pursuit of their common and individual interests. Participation by successful bidders is entirely voluntary, as detailed in a partnering paragraph in the solicitation for bids and calls for sharing by the two parties of costs associated with effectuating the partnering arrangement.

The Mobile District of the Corps, in its *Guide to Partnering for Construction Projects* (*13*), sets forth practical, straightforward suggestions for the implementation of project partnering. Information is provided on the following items, among others:

- Getting a timely start.
- Obtaining top management commitment.
- Identification of a sponsor or "champion."
- Selection of team members.
- Selection of facilitators.
- Workshops and follow-up actions.
- What to expect in the way of special costs.

The Mobile District guidelines also address the application of the partnering concept to smaller projects for which the scope and complexity would not warrant the full treatment accorded the large, complex projects on which the concept

has been successfully piloted. The central ingredients are mutual trust and cooperative attitudes, with a minimal requirement for formal team-building workshops, and so on.

Whereas the Blue Ribbon Panel's recommendations and the more recent initiatives focus on responsibilities for actions within the Corps of Engineers, partnering is seen as the prime mechanism for enlisting the participation of the contractor in a quality-intensive effort. The Corps has documented significant improvements on three pilot applications: Oliver Lock and Dam Replacement, Mobile District; NASA Test Operation Control Center, Mobile District; and Bonneville Navigation Lock Project, Portland District. The improvements have been in the areas of safety, cost control, schedule improvement, value engineering savings, dispute avoidance and resolution, and overall in improved attitudes of mutual respect and enthusiasm among the key stakeholder-participants in the projects. The Corps is considering some pilot applications of partnering with AE contractors and the inclusion of the non-federal cost-sharing sponsors for which it performs civil works projects.

Naval Facilities Engineering Command (NAVFAC)

NAVFAC also has begun a commandwide, systematic effort to instill TQM philosophies and techniques using a top-down approach. For example, NAVFAC's Southern Division, headquartered in Charleston, South Carolina, is concentrating on creating the necessary TQM environment, with an Executive Steering Group selecting a set of key processes for subsequent team improvement activities. Initial quality training is well under way, concentrated at the management and facilitator levels. Function-specific training also is receiving new emphasis. Appropriate use is being made of technology systems to reduce errors, variation, and cycle time. NAVFAC also recognizes exemplary performance by design firms through its Design Excellence Award system. Any firm whose performance on a single project is considered noteworthy with regard to the following criteria can be given the award:

- Mission support,
- Cost-effectiveness,
- Low change order rate,
- Appropriate design,
- Quality control,
- Timely design effort, and
- Cooperation.

Appropriate consideration is given to architect-engineer firms that have received a Design Excellence Award when they apply for subsequent projects (*14*).

BEYOND THE AGENCY LEVEL

Many examples doubtless exist of other successful TQM implementations by public agencies, possibly agencies involved in design and construction. It will require a number of success stories such as the Oliver Lock and Dam Replacement project to stimulate governmentwide adoption of TQM as the core of all management methods. One indication that interest in TQM is spreading beyond the Executive Branch of the federal government is a report (*15*) recently prepared for Congressman Donald Ritter by the General Accounting Office (GAO). The GAO study reviewed the TQM achievements of 22 companies that were among the highest-scoring applicants for the Malcolm Baldrige National Quality Award. Among the major findings were that these companies experienced better employee relations, improved quality with lower costs, greater customer satisfaction, and improved market share and profitability. It is not known what use will be made of the Ritter Report, but it creates a degree of hope that TQM has "infiltrated" another sector of government and that continued deployment of knowledge of TQM's benefits will eventually lead to formal replacement of some of the quality-inhibiting practices in government, including public construction.

CONCLUSION

Can TQM work in public construction? The answer, based on the examples and discussion presented in this paper, seems clearly to be "Yes, it can." Despite widespread popular opinion, there are no insurmountable structural deterrents in the current system of contract procurement that rule out a fairly comprehensive implementation of TQM principles and methods on a single project, on a succession of projects such as an annual appropriation program, or within and between the contracting parties. It is true that such effects as short-term rotation of top managers, quality-inhibiting personnel reward systems, and restrictive procurement regulations make it more difficult to go all out on TQM implementation in public construction than in the private sector. However, private contractors and owners have stated to us that they encounter equally frustrating implementation problems—involving the civilian version of some of the same issues—in initiating and sustaining TQM processes in their side of the industry.

The more pertinent question is, "Will TQM work?" The answer lies in the common element between the private and the public sectors, whether it be in construction or in manufacturing, and that is people. The TQM process, to use an over-used term, is not "rocket science." It has more accurately been described as "organized common sense." The success of TQM in public construction will depend on the knowledge, the dedication, and the perseverance of the people, both contractor and agency, who come together contractually to execute a project. Moreover, it is the people at the top of each contracting party's organization on whom the success of the joint implementation squarely rests. The respective and joint leadership teams must understand and believe in the principles and elements discussed above. Their belief must be akin to passion for the process to have a chance and for the numerous other people in the process to believe in the TQM approach and be able to effectively contribute to it.

As for the restrictive personnel and procurement practices, they also are the product of processes put in place, and fully amendable, by people. If those agencies, agency managers, civil servants, military personnel, contractor managers, and contractor personnel who are now engaged in the innovative, "envelope expanding," projects described in this paper will persevere in the effort and allow themselves to adopt TQM work principles in an almost religious manner, then the old

paradigms will continue to be replaced and the line of would-be participants in TQM-based public construction will become longer. In time, the writers of the restrictive policies may themselves be sufficiently influenced by the new art-science to revise the regulations to reflect more of a TQM-friendly environment.

We believe that government and its private contractors bear approximately equal shares of the burden for making TQM work in public construction. However, the "owner" must always take the lead by setting the necessary environment for cooperative contracting and then enabling the process to work by establishing proper processes for quality-related cost and risk sharing, routine project management, and dispute avoidance and resolution. If this basis is provided by the owner, as it is in the Corps of Engineers and FHWA partnering systems, then the burden shifts to the contractor to (a) believe that TQM and public construction can coexist, (b) have a serious implementation of TQM in place internally, and (c) be truly motivated to TQM-based performance on the project at hand.

With the public agencies taking the lead and the design and construction contractors accepting the challenge, TQM indeed can and will work in public construction.

ACKNOWLEDGMENTS

We express our gratitude to the following persons for their cooperation in providing information for this paper: Thomas Edick, Federal Highway Administration; Mark Grammer, U.S. Army Corps of Engineers; and Dale Johannesmeyer, Southern Division, Naval Facilities Engineering Command, Charleston, S.C.

REFERENCES

1. L. Dobyns. *NBC White Paper: If Japan Can, Why Can't We?* National Broadcasting Company, 1980.
2. W. E. Deming. *Out of the Crisis*. Massachusetts Institute of Technology, Cambridge, Mass., 1986.
3. Roadmap to the 90s: Achieving Excellence in the Public Service. In *Proceedings of the Third Annual Conference on Federal Quality and Productivity Improvement*, Office of Management and Budget and the President's Council on Management Improvement, 1990.
4. D. S. Holt. Development of Quality Management in Minnesota. *Hot Mix Asphalt Technology*, Winter 1990, pp. 11–12.
5. P. Hughes. Mn/DOT's QIP Program. *Hot Mix Asphalt Technology*, Winter 1990, pp. 12–13.
6. G. E. P. Box, L. W. Joiner, S. Rohan, and F. J. Sensenbrenner. Quality in the Community: One City's Experience. *Quality Progress*, Vol. 24, No. 5, May 1991, pp. 57–63.
7. M. Walton. *The Deming Management Method*. The Putnam Publishing Group, New York, 1986.
8. M. Walton. *Deming Management at Work*. G.P. Putnam's Sons, New York, 1990.
9. J. L. Burati, Jr., and M. F. Matthews. *Quality Management Organizations and Techniques*. Source Document 51. The Construction Industry Institute, Austin, Tex., 1989.
10. *Total Quality Management: The Competitive Edge*. Publication 10–4. The Construction Industry Institute, Austin, Tex., 1989.
11. T. Shoop. Uphill Climb to Quality. *Government Executive*, March 1991, pp.17–19.
12. *Final Report: Implementation of the Blue Ribbon Panel Recommendations on Management of Construction Quality in the U.S. Army Corps of Engineers*. Office of the Chief of Engineers, March 1987.
13. *A Guide to Partnering for Construction Projects*. U.S. Army Corps of Engineers, Mobile District, Jan. 1990.
14. Design Excellence Award Program. *CEC Biweekly*. Naval Facilities Engineering Command, April 1989.
15. *Management Practices: U.S. Companies Improve Performance Through Quality Efforts*. Report GAO/NSIAD-91-190. General Accounting Office, Washington, D.C., 1991.

Quality Management for Concrete Pavement Under Performance Standards

JAMES M. SHILSTONE, SR.

Alternative methods are being researched to develop ways to replace traditional highway contracting practices with performance standards. This will force major changes in both agency and contractor responsibilities for construction planning and quality management. Engineering formulas can be used to calculate structural solutions. These formulas do not resolve durability needs. There is no accepted means to define, in measurable terms, the characteristics of a concrete pavement that will be durable in typical highway environments. Methods are described whereby existing pavements with excellent and poor performance histories may be surveyed and used to define "durable concrete" and act as the basis for performance standards for durability. The methods for specifying, controlling, and verifying construction to meet the desired durability objectives are outlined.

The term "quality" is often discussed, universally supported, and everyone's acclaimed policy. "Quality" has become an idealized subject like motherhood, clean water, and zero pollution.

However, the Business Roundtable Construction Industry Cost Effectiveness Project Committee report opened with the statement, "By common consensus and every available measure, the United States no longer gets its money's worth in construction, the nation's largest industry." This appraisal was later echoed in the National Research Council report on the highway system (*1*). These findings reflect the failure of the traditional design-construction-inspection process to produce quality projects.

Two of the most respected leaders in the quality movement are W. Edwards Deming and Philip Crosby, but each defines quality in different terms. Deming describes it as "customer satisfaction," whereas Crosby refers to "conformance." The objectives are identical but each speaks to a different point in the chain of raw materials to the consumer. Deming speaks of the gap from manufacturer-supplier to buyer. Crosby addresses the subject from the worker with his tools to the product as designed.

Although zero defects (and assurance of quality) is a good objective, it is not realistic. During the 1983 National Conference on Quality Assurance in the Building Community, Geoffrey Fhronsdorff of the National Institute of Science and Technology described the Battelle Memorial Institute report of early product failures described as "fractures." The objective was to develop a correlation between "cost of prevention" and "cost of fracture."

It was reported that, for a nominal expenditure, 28 percent of unacceptable levels of quality can be prevented. For a higher cost, an additional 35 percent can be prevented. The remaining 37 percent can seldom be prevented regardless of the expenditure. Current state-of-the-art design and construction are not sufficiently advanced to cope with all problems, human errors, and technical factors that do not interface properly.

Shilstone & Associates, Inc., 8577 Manderville, Dallas, Tex. 75231.

One problem is the terminology. The expression "quality assurance" creates a false impression of what is to be provided. By definition "assurance" means "the act of assuring; the state of being sure or certain." The only factors of which we can be assured are death and taxes. Even the high-level quality assurance program used in the U.S. space industry failed for one of the missions. How can the construction industry expect better success using its traditional practices?

A rational alternative term is "quality management." The responsibilities for management are well understood: someone is responsible for supervising and directing the work. "Quality management" is defined as all systems and methods needed to produce a product that will meet the assigned criteria with a predetermined degree of reliability. For highways that responsibility starts with the design engineer. A contractor cannot cast a high-quality pavement following a poorly conceived design or antiquated specifications.

Darrell W. Harp, Assistant Commissioner for the Office of Legal Affairs of the New York State Department of Transportation, said of this system:

> The contractor can't use his own initiative because he has little option when he is told precisely what he must do, what type of materials he must incorporate and exactly how it is to be put in place. Innovation is stymied. Another drawback is that the improvement of the product will be very slight and it is doubtful that you will see a reduction in overall cost. If we were to live forever with materials and methods specifications, I suppose we would still be driving around in Model 'Ts'. (*2*)

DEFINING THE PROBLEM

During the In Search of Excellence Conference in Hawaii, Damian Kulash identified three steps to quality:

1. Define what we mean,
2. Choose effective ways, and
3. Convince everyone of what is wanted and how to do it.

Once the first step has been accomplished, the second step can follow. Solutions to the third step will be difficult because it involves a paradigm shift requiring that old ways give way to the new. If a different result is wanted, something different must be done. The change will not be as difficult as some

might believe, but all change is frightening and met with opposition. That is a problem of major proportions for the highway industry. One of the hurdles will be political leverage that may not be in the best interests of the driving, taxpaying public.

This paper provides a solution to the first two steps. Yogi Berra put Kulash's first step in a simpler way when he said, "If you don't know where you're goin', you might end up some place else." Despite decades of highway construction, there is no means to quantify, in technical terms, the characteristics of a durable concrete pavement. Millions of dollars are being spent to figure out how to design something that cannot be defined. Good design formulas do not ensure that a durable pavement will be put in place. More knowledge of how materials interface and how to cast a pavement to produce a dense, durable composite is needed rather than more design formulas.

In FHWA Publication FHWA/RD-87/095 on the relationship between concrete consolidation and performance, the summary includes the following statement:

> A concrete pavement is a manufactured item. Therefore, it is necessary to ensure that quality control is exercised at all critical phases of production. Presently, quality control specifications exist to assure a quality product through all phases of construction up to delivery of concrete at a site. However, there are no direct specifications available to monitor concrete quality after the concrete is placed in front of the paver. (*3*)

Quality of the constructed project, in terms of both strength for loads and durability for the environment, must be defined in measurable terms. Engineering design formulas provide solutions for structural needs but not for durability. Concrete pavements that meet design criteria for thickness and strength may not be durable or produce a smooth ride. Therefore, good design must include both good engineering design and sound technology.

Today construction is controlled on the basis of input with the expectation that output will be consistent. The paving industry buys sand, stone, cement, asphalt, steel, and other products that have broad producer-oriented tolerances. The result is cumulatively variable based on the effects of the tolerances. It is like expecting the sum of 16 + 22 + 12 to produce a total of 50. Actually the sum can vary between 41 and 59 when each of the factors is allowed to function within variables such as $16\,(\pm 3) + 22\,(\pm 4) + 12\,(\pm 2)$. This is what happens with most building materials. Acceptance of the broad tolerances allowed for input items ensures variable output.

The quality management concept suggests that control be based on recognition of output objectives and that adjustments be made as materials vary during production to deliver the desired quality. In simpler terms, become performance oriented and manage the production to produce the desired results. This was the policy followed by highway department project engineers years ago when bidding was by line item. The engineer monitored progress and directed changes in midcourse to produce the intended result.

DEFINING THE QUALITY STANDARD

The quality process involves determining the desired qualities of the constructed project, translating those qualities into measurable terms, and managing the work to produce those qualities.

Like beauty, quality is measured in the eyes of the beholder. Three major groups are interested in the quality decision: the public, the agency-engineer-inspector, and the contractor and his suppliers. Each looks upon quality in different terms. Once the quality objectives of each are understood, a plan can be developed to pay the bill for and produce the needed quality.

- The public wants personal comfort, minimal interference with their lives, and low cost. These do not translate to mathematical formulas. Although low initial cost is always of concern, long-term durability is also important. If the work is not durable, the repairs interfere with traffic and raise taxes. Therefore, the engineer-agency's customer is the driving public who wants a better ride, flat pavements, and no roller coasters. The public measures quality by comfort and durability—that is Deming's "customer satisfaction."
- The engineer is responsible for planning the public's measure of quality. The problem is that an engineer thinks in terms of mathematical formulas. On the basis of tests and theories, the engineer calculates loads and loading cycles and comes up with a solution to satisfy the public. Then bid documents are prepared that describe everything and how to do it in engineering terms. Often high, costly safety factors are used to overcome what is perceived to be poor potential execution of the work. Physical tests are made to verify only a minimal level of ingredient quality. To the engineer, quality is measured by numbers.
- Finally, contractors and suppliers must follow Crosby's objective for conformance but are forced by the bidding process to make price-oriented decisions. They bid the least cost to meet the minimum requirements of the contract. If they bid to provide solutions that meet the engineer's and the public's intent, they are not competitive. Therefore, the construction team measures quality by conformance at the least cost.

Failure in the quality system derives from these divergent means of measuring quality. The conflict is among the public's desire for personal comfort, the engineer's love for numbers, and the contractor's concern for dollars. The actual bottom line is dollars, although that does not always contribute to customer satisfaction for the taxpaying public.

If there are problems on a project and resolution is through litigation, the contractor generally prevails. Contractors are protected by the 1918 U.S. Supreme Court decision *U.S. versus Spearin*, under which an owner warrants that the plans and specifications will achieve the results intended. The contractor is not bound to use the most desirable of all the options and can do anything within the limits of the contract. If the contract specifies tolerances that can be combined to produce an unacceptable product, the owner—not the contractor—is responsible.

Until recently highway departments, and civil engineers in general, worked by line item for bidding and construction. They controlled quality by making adjustments in the field to meet their objectives. That system has been changed to one of hard dollar bids. Despite the change, the department of transportation manuals are essentially the same for materials and provide broad, often politically motivated tolerances. The

construction engineer no longer has the power to require adjustments to control the process and guide the work to produce the intended quality. The contractor can use the worst of all combinations, and the public is forced to accept the results. The fault does not lie with the contractor, who is forced by the low bid practice to survive in the price jungle.

INNOVATIVE CONTRACTING AND PERFORMANCE STANDARDS

TRB's Task Force on Innovative Contracting Practices has been evaluating ways to overcome past problems and meet the needs of the future. One of the major problems facing the agencies is the reduction in personnel. Figure 1 is an FHWA graphic representing the projected relationship between federally funded projects and available agency personnel based upon assumptions that the 1991 highway act would be passed as originally planned. Clearly shortages in personnel will make it impossible in the future for agencies to control work in the same manner as they do today.

There is no question that agencies are losing many of their most experienced personnel. Most replacements have had little training in basics, much less innovations in materials technology. A state concrete engineer who was to retire within 2 years said, "You must remember I am a product of my agency. I have worked here 30 years and been taught only to follow the Manual. I don't know any technology other than what I was told by the department." An interested, younger man was assigned as his replacement but was reassigned after several years because he was performing well and he could fill an immediate need. There is no qualified replacement for that concrete engineer, although he will retire this year. How will someone be able to take his place?

Many officials also admit that a major objective is to develop a program to prevent agencies from losing so many claims. Obviously alternative contracting and quality management practices must instituted.

DEFINING OBJECTIVES

It will not be a complex task to determine the characteristics of a good concrete pavement. There is no need for laboratory research to reinvent good concrete. Hundreds of miles of highways are performing extremely well. Some have been in place up to 100 years. What is known about why they have done well? There are also hundreds of miles that are deteriorating at a faster rate than that expected. These two extremes are ideal research specimens. They have been affected by the real-world conditions. No accelerated, artificial laboratory tests are needed. The in-place construction is there to be examined for clues as to what made each perform as it did.

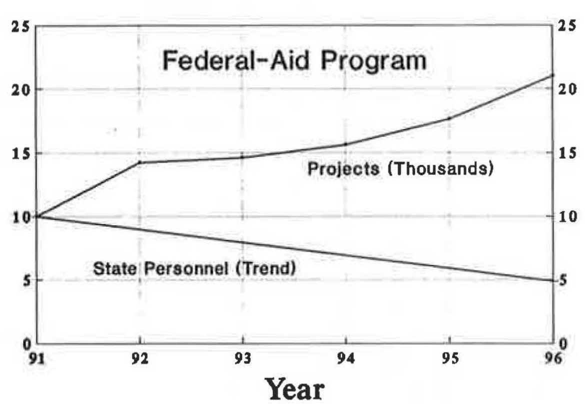

FIGURE 1 FHWA projected relationship of federally funded projects and available public agency personnel under the proposed 1991 highway bill.

The investigators should be technologists who understand the concrete process. They need know little about engineering design formulas but much about concrete-making materials and the concrete construction process. They will have to correlate the effects of materials, mixtures, placement practices, consolidation, curing, and testing with performance. Textbooks do not explain the problem. The investigators must have inquisitive minds and be ready to report that some of the highly respected books written 40 years ago do not necessarily relate to today's materials and needs.

Concrete forensic engineers with whom I have discussed this approach say that it is innovative. The senior petrographer for a major federal agency said that in his many years of work with that agency, he had been asked on only one occasion to find out why a certain concrete performed well, yet he had studied thousands of failures. The profession is conscious of failures but takes little time to study and accentuate quality.

The study program should proceed as described in the following sections.

Identification of Subject Pavements and Background Material

Identify old and new, good and bad pavements and assemble records about each. If available, include mixture proportions, adjustments, mixer types, placement methods, consolidation equipment and methods, reinforcing details, joint details, finishing methods, curing, and climatological conditions at the time of casting.

Specific information—not just reference to compliance with a standard—is needed. Contract documents may be helpful, but they do not provide the detail needed. Assumptions will have to be made for very old projects. Retired personnel and old documents might provide needed background. Investigators must persevere in their search for project information and the practices of the time.

Surface Study

Perform a surface study of the selected projects and observe and record overall and detailed conditions. Identify deterioration that may be based upon engineering design, excessive loading cycles, or base failures. These should not be allowed to interfere with the study of the concrete portion.

The study should reflect the occurrence of scaling, D-cracking, carbonation cracking, aggregate reactivity, abrasion, erosion, or other concrete deterioration. Special attention should be given to variations in measurable qualities of work done under the same contract. Investigators must be prepared to ask such

questions as whether concrete permeability was high or low where D-cracking occurred. This can possibly be answered by finding portions of a pavement where the same materials were used but one did not show D-cracking.

Photograph area and typical surface conditions. The surface details should be shown within an 8-in. (±) square framed by a grid cut in grey card stock with ⅜-in. hatch marks around the edge. Figure 2 shows the setup used in a study for the city of Dallas.

Figure 3 is a closeup of a 1922 cast pavement still in service. It has a high density on the surface of particles passing the ⅜-in. and retained on the No. 16 sieve. That aggregate distribution appears to be in conformance with ASTM C 33 and the Portland Cement Association (PCA) recommendations of the period. The 1923 issue of ASTM C 33 required that the sand consist primarily of coarse particles. *Design and Control of Concrete,* Issue 1, recommended that only 65 percent of the fine aggregate be allowed to pass the No. 8 sieve.

The American Concrete Institute reports from Committee 201, Guide to Durable Concrete (*4*), and 210, Erosion of Concrete (*5*), cite the need for the intermediate aggregate particles to minimize abrasion and erosion. Intermediate-sized particles have other beneficial effects on the concrete, such as improvement in rheology during placement, resulting in increased density and decreased permeability.

Figure 4 shows a new pavement that was wearing rapidly. The gap between the ½-in. particles and the mortar is distinct. The aggregate gradation is acceptable under current standards. The 1989 ASTM C 33 gradations provide for coarser coarse aggregate and finer sand. Gap grading of aggregates appears to contribute to durability problems in many ways.

Correlation of Data and Definition of Pattern

Correlate data and photographs from the projects studied and attempt to define a pattern from the study. In most cases, trends will be found during the detailed examination. Data can help identify the causes of performance differences. From that study, representative locations should be selected for physical testing and petrographic study.

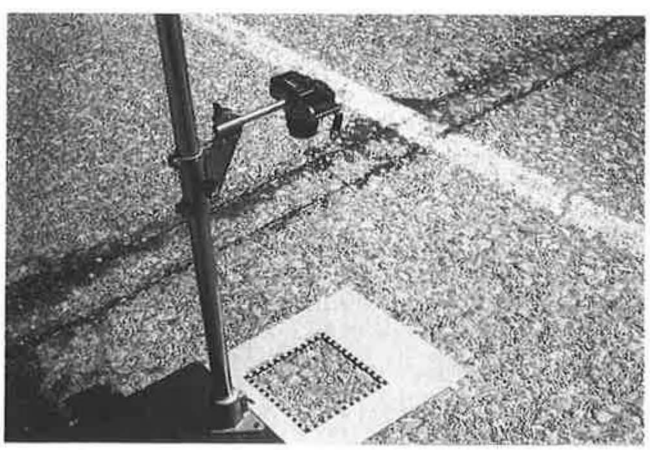

FIGURE 2 Photographic setup for Dallas study.

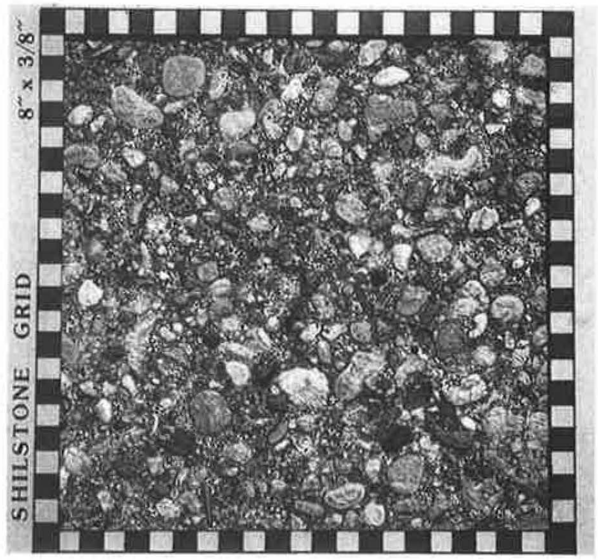

FIGURE 3 1922 cast concrete pavement with high aggregate density.

Physical Testing

At each selected sample site, drill parallel holes and use a nuclear density gauge similar to Troxler 2376 (see Figure 5) to measure variations in pavement density from the bottom to the top at 1-in. intervals. Cut 8-in.-diameter cores between the drilled holes to include the concrete evaluated by the nuclear gauge.

Take sufficient 4-in. or 6-in. cores to make physical tests, tests of permeability by the boil or chloride penetration method, and petrographic studies.

Examine the base under the core location to assess its potential effect upon the concrete condition.

Cut the 8-in. core into 1-in. (±) plates and evaluate each plate petrographically and by other means to determine variations in permeability, relative density, and trapped and entrained air.

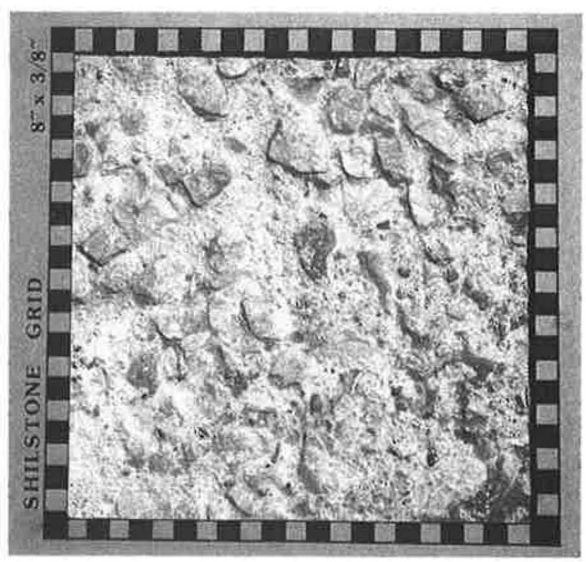

FIGURE 4 Poorly performing gap-graded mixture.

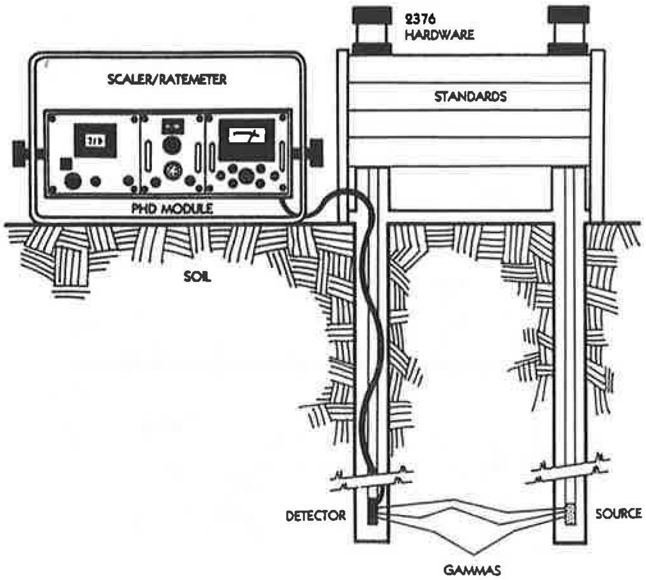

FIGURE 5 Nuclear gauge to measure variations in relative density.

Determine compressive strength, split core strength, and permeability for the other cores. Figure 6 is a regression analysis comparing permeability and compressive strength of companion cores from the Dallas study. The pavements with data in the upper left quadrant are performing well. Those in the lower right quadrant are performing poorly. Those in the other two quadrants have special features to overcome deficiencies in one of the measured results.

Petrographic Studies

Perform petrographic studies of some of the cores, not only to assess the quality of the paste and mortar, but also to identify reactive aggregates that may have contributed to problems. Describe the aggregate particle shape, texture, and distribution by individual sieve sizes to determine the combined gradation. Describe the aggregate distribution by particle size, not just as "coarse" and "fine." The extremes will be (a) those that are well graded with dense distribution and (b) those with a high incidence of one or two coarse aggregate particle sizes and fine mortar.

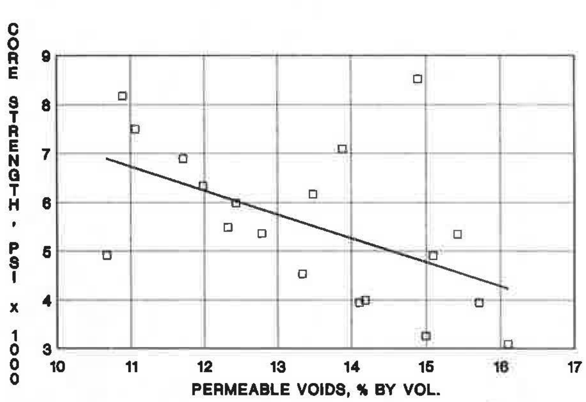

FIGURE 6 Regression analysis: permeability versus core strength.

Analysis and Report

Analyze and report the data. The report should provide a picture of the difference between the good and the poorly performing pavements. It is anticipated that the following factors will be found to influence performance:

1. Relative density difference between the top and bottom of the cores, with the lower density on the bottom. It is anticipated that uniform, high-density concrete will perform better.
2. Permeability will play a significant part in the durability, especially where water penetration has contributed to D-cracking or sulfate problems.
3. The balance between large (trapped) and small (entrained) air will be influenced by aggregate shape and the amount of coarse aggregate in the mixture.
4. Freeze-thaw resistance will be improved by both the relative density of the concrete and the type of air. It is expected that pavement with lower than normally specified air content will perform well. This will be accentuated in concrete cast before air entrainment was used. A lesser amount of purposely entrained air with a low spacing factor will assure better durability than will a large amount of trapped air.
5. Well-graded mixtures perform better than gap-graded mixtures. The high incidence of ⅜-in. to No. 16 particles helps block capillary pores and improve abrasion resistance.

PERFORMANCE SPECIFICATION

On the basis of the research, specifications should be developed that describe the performance required (*6*). They should be brief and results oriented and avoid defining how the contractor should do the work. Current state manuals should not be referenced. It is not possible to change results while keeping the recipe the same.

The contractor's responsibility for quality management must be defined. Only the organization performing the work can control the quality. That function must include in-process controls and in-place verification of hardened concrete quality. The contractor should be required to plan and organize the construction process using technical skills to assemble materials, equipment, and labor in compliance with the performance standards. This planning will ensure that primary attention is given to results, rather than resources.

The contractor should not have to labor under the constraints of prescriptive requirements but should have the maximum latitude to produce the pavement at the lowest cost.

The following performance objectives can be used as a guide for details to be covered in the performance specification:

Engineering: strength, thickness, and jointing;
Construction: compacted density, surface texture, and ride;

Durability: entrained air; sulfate resistance, D-crack controls, or both; abrasion or erosion resistance, or both; and permeability limits;

Quality: responsibilities, planning, materials, equipment, methods, personnel, and control practices;

Other: incentives and disincentives;

Verification: owner methods.

The tendency to revert to traditional methods must be avoided. The fewer words, the better. If the document describes "how," it is not a performance document. It must describe "what" as an end product, but not its components. For example, the qualities of potentially reactive aggregates can be described for durability, but stockpile gradation must defer to combined gradation in the mixture if that is found to be a key to performance.

Incentives and disincentives are important for success. The past practice has been to assess penalties and give no credit for higher quality, even though that quality may extend pavement life. There are many examples in which the lack of provision for incentives costs the taxpayer benefits. The following is an example.

A new technology allowed a contractor to increase flexural strength by more than 150 psi. Since he was paid for only the lower strength and a cement factor was not specified, he reduced the cement content by 50 lb to save about $1.50/yd². The added strength could have extended the life of the pavement because of the lower water-cement ratio. The higher quality was of more importance than the $1.50/yd² that the contractor saved. The taxpayer is the loser.

Some might respond to this by advocating the specification of a high cement factor, "to be safe." This can be counterproductive. The contractor has no incentive to produce a strength higher than that specified. He can use more sand (which is often cheaper) and still meet strength. High-sand mixtures require more water. Durability is influenced more by total water than by the water-cement ratio. In the end, durability and the taxpayer suffer.

CONTRACTOR QUALITY CONTROL

The president of an international construction firm contacted Shilstone & Associates on the subject of quality while bidding on a project in Saudi Arabia. He said that he wanted a program set up on the job so that when the superintendent wanted to do something under a tight schedule, the correct materials, equipment, and trained personnel would be there to "do it right" the first time.

The contractor is, as is any manufacturer, the only one who can control the quality of his work. If he is to have responsibility for furnishing a product, he needs the authority to control how it is done. In the late 1800s, John Ruskin said, "Quality is never an accident. It is always the result of intelligent effort." The "intelligent effort" starts with an approved quality control program and qualified people assigned to perform the work. Deming agrees with that concept and said, "Quality must be built into the work. It cannot be inspected in."

The first step in quality management is to develop a plan of action. Most planning starts with the resources—the baseline items. Quality planning works in reverse. It starts with the constructed product and the question, "What are the factors that affect attainment of the primary objective and those that will produce incentive pay?" This forces the planner to consider all of the variables and how to best mesh them to meet his objective. The principal factors are materials, equipment, methods, and people. Books have been written on the subject, but the following are a few of the major factors.

Materials include not only mix ingredients, but also the methods of curing, joint sealants, and other products. Not all cements are equal. It must be known which will produce not only the long-term strength, but also the early strength to allow access to completed castings. The same is true for admixtures and fly ashes. The chemically active ingredients should be evaluated to determine their chemical compatibility. The admixture suppliers should be asked if a special sequence of addition should be followed.

Aggregate distribution has a major effect on water demand. A well-graded mix requires less water and is more responsive to vibrators and finishing. Since only the combined gradation is to be considered, the most economical, locally available materials can be combined without the need for stockpiles to meet arbitrary gradation requirements.

Colorado Department of Transportation engineers replaced 500 lb of ¾-in. gravel and 300 lb of sand with 800 lb of waste "squeegee" (⅜ in. to No. 30) aggregate (*7*). This reduced the water by 5 percent, increased the strength by 10 percent, and changed a difficult-to-finish bridge-deck mix into one that "finished like butter." A materials supplier in another state reported wasting 1,000 tons per day of nonstandard-gradation aggregate. This material now accounts for 32 percent of the combined aggregate for an excellent mix.

Occasionally a cheap sand may be acceptable but will require more mix water and produce an acceptable but lower strength. With an incentive clause in mind, the quality manager may find that a more expensive coarse sand or blend of other materials may be more expensive, but it can

1. Reduce water and produce higher strength,
2. Increase relative density,
3. Improve mix mobility,
4. Make finishing easier,
5. Produce a better ride, and
6. Ensure incentive pay.

Equipment includes that for batching, mixing, placing, compacting, finishing, and testing. The more bins or means to use multiple aggregates, the better. The United States is the only industrial nation that allows use of a single coarse aggregate and a single fine aggregate. Such a system makes it impossible to blend materials to improve mix characteristics. This can be partially rectified by use of cold feed systems (as used for asphalt) to blend aggregates before placing the mixture on the hopper loading belt. In my opinion, a minimum of three aggregate bins is necessary to manage mixture proportions, but four are helpful.

Mixer quality affects strength and mixture workability. Placing, compacting, and finishing equipment must be compatible with the mixture, because it is key to producing high relative density and low permeability. The desired degree of consolidation must be uniform throughout the casting. A mix-

ture that responds well can be compacted with good, but lighter equipment working at a higher speed to improve productivity. The contractor must demonstrate that the performance standards will be met by casting and testing a trial section before start of the work.

Testing equipment must be appropriate, calibrated, and in working order. The FHWA demonstration program is reporting that modern testing procedures, including maturity meters and ultrasonic equipment, provide more information than the traditional physical tests. Questions about entrained versus trapped air content are being heard. The conventional pressure and volumetric test equipment must be in good working order. Wet unit weight should always be checked, because it can identify problems with air meters and verify yield. The physical testing equipment and test practices should be in conformance with ASTM standards.

Methods used must be compatible with the objectives. It is anticipated that the recommended investigation will uncover some current methods that do not work. Even the previously quoted FHWA report on consolidation (*3*) indicates that the industry has little quality control over concrete once it has been placed ahead of the spreader. Performance standards will make it essential that the methods used work. If they do not, the disincentives will be costly for the contractor.

Tests should be meaningful. Although the slump test has been used for many years, it describes only the consistency of a batch and has little or no relationship to strength potential. The wet unit weight in a $\frac{1}{2}$-ft^3 bucket is the best all-around test. The weight is affected by both water and air content.

The methods for casting test specimens are important. ASTM Standard C 31 provides that specimens cast at a slump of 1 in. or less must be consolidated by vibrator. Slumps from 1 to 3 in. may be consolidated by a vibrator or a rod. Vibrated specimens are more representative of in-place concrete placed using modern paving equipment.

The effects of vibration versus rodding can be significant. To confirm this and the applicability of vibration at any slump, a series of 50 sets of companion cylinders was field cast using both puddling and vibrating methods. The slumps ranged to 6½ in. A statistical analysis of the data revealed that slump was immaterial. The compressive strength of vibrated cylinders averaged 450 psi higher than that for rodded cylinders.

Personnel is critical to success. Many of the most experienced agency and construction supervisors are retiring. There has been little effort to train others to fill the technical gap. In terms of quality management, the loss of the technical skills is more important than loss of operational experience.

Technology is the engine that will drive quality in the future. Instead of an increasing number of personnel in the technical pipeline, the number is shrinking and the pipeline is almost closed. Most engineering universities are almost entirely eliminating the study of materials from their curricula. Who is teaching the future leadership? Who is teaching advanced concrete technology? Most technologists have gained their experience in the school of hard knocks. Few of the young, more skilled quality control personnel stay with that part of the industry because the pay is better in other segments.

The smart contractor of the future will apply not only operational, but also technical expertise to maintain a strong position. He will apply the correct technology to do it better at a lower cost and also get full incentive pay. His problem lies with the lack of trained personnel to fulfill the need and the few places where his own personnel can go to be trained in the practical applications of construction technology. The American Concrete Institute technician program is good, but it doesn't go far enough.

Field personnel must be trained and not just hired to do a limited job. Video provides an excellent opportunity for training. Even the economical Camcorder can produce an adequate training tool. From experience, I have great faith in the American worker. Given an opportunity to learn and improve his position, he will do so. Lend him a video tape to help him better understand and hold his job, and he will find a way to watch it.

The less formal the presentation, the better. It should not be a military-style presentation or a "rah-rah-we-are-a-great-company" motivator. It should be a personal, one-on-one, factual description of what the workers can expect. Show them what makes a pavement good and what they can expect to see. Help them identify tell-tale signs and teach them how to do their job. Learning from others who probably do not have the right answers is not the way to get started.

Ask the workers their opinions. That process, known as Quality Circles, has been extremely effective in Japan and the United States. Feedback from those doing the work is a more effective means to improve the quality of the work than an idea generated by someone far from the work. Books have been written on the subject, and the American Society for Quality Circles has done much to advance the applications.

Rewards and incentives to workers should be part of the quality program. Some rewards should be published and others, spontaneous. If a worker knows that management can receive incentive pay, but he gets none of it, he will respond as do the contractors who do not have incentive clauses. Yes, a share of the benefits in dollars is important, but not the only thing workers seek. According to many studies of worker psychology, nonmonetary personal rewards are high motivators. A Gold Hat Club could signify that the wearers have been recognized for their quality work. There should not be many of those, because it would weaken the image. An unexpected barbecue and beer supper halfway through the job will create a better attitude. The best and easiest motivator is a pat on the back and a thumbs-up sign. Good human relations improves quality and should be an integral part of the quality system.

AGENCY RESPONSIBILITIES

Once work gets under way, the agency must audit quality control. Acceptance testing should be done during construction and after completion. Performance contracts require the agency to undertake their work from a new point of view. They will have to review submittals and judge their adequacy to meet both short- and long-term objectives. Performance warranties probably will not exceed 5 years, although the objective may be 40-year durability.

Those who review submittals will not be able to rely on formulas similar to those used for design. Their concern must be durability, and be based upon technical factors. For example, a given combination of materials may produce ade-

quate concrete strength for load-carrying capability for the pavement. However, that combination may contribute to long-term problems due to poor rheological properties and experience rapid wear in high-traffic or turning lanes. Those reviewing the documents must be knowledgeable in construction and not those who follow a checklist based on manual limits.

CONCLUSIONS

There is no simple method to implement this important subject. Quality is the product of hard, properly directed work. This program offers a new venture into the unknown but one that can result in construction that meets the needs both now and in the long term. The benefits of a properly established and managed quality program are clear:

1. The objective, proven from 50 years of testing and field experience, will be defined, targeted as a goal, the center of the planning process, and verified by measurement.
2. Litigation will be reduced.
3. Innovation will be expanded to increase productivity, improve quality, and reduce costs.
4. Higher-quality pavement will be better able to span base materials problems.
5. Quality in the form of customer satisfaction will be provided for the taxpaying public.

REFERENCES

1. Committee on Concrete Durability: Needs and Opportunities. *Concrete Durability: A Multibillion-Dollar Opportunity*. National Academy Press, Washington, D.C., 1987.
2. D. W. Harp. Innovative Contracting Practices: The New Way to Undertake Public Works Projects. *Hot Mix Asphalt Technology*, Winter 1990.
3. Concrete Technology Laboratories. *Relationship of Consolidation to Performance of Concrete Pavements*. Report FHWA/RD-87/095. FHWA, U.S. Department of Transportation, 1987.
4. Guide to Durable Concrete. In *ACI Manual of Concrete Practice*, Part 1. ACI 201.2R-77. American Concrete Institute, Detroit, Mich., annual.
5. Erosion of Concrete. In *ACI Manual of Concrete Practice*, Part 1. ACI 210R-87. American Concrete Institute, Detroit, Mich., annual.
6. J. M. Shilstone, Sr. Performance Specifications for Concrete Pavements. *Concrete International*, Dec. 1991.
7. The Importance of Gradation: A Tale of Two Mixes. *Concrete Construction*, June 1991.

FHWA Demonstration Project No. 89 Quality Management and a National Quality Initiative

DONALD R. TUGGLE

Quality assurance specifications and programs in the highway construction industry have been evolving since the 1960s. Within the last decade there has been increasing attention to promoting quality products and services throughout the U.S. economy. There has also been an increased level of interest within the highway community. Although there is currently significant interest and many independent activities associated with what has now become known as quality management, there is a need to coordinate these many activities. There is also a need to increase awareness in and build support from upper management, and to provide technical skills and tools to those responsible for implementing quality management programs and specifications. A coordinated effort among the Federal Highway Administration, the American Association of State Highway and Transportation Officials, the highway construction industry, and others is being formulated to provide oversight and direction toward increasing emphasis in quality management and other construction quality and performance issues. This effort has been termed the National Quality Initiative. FHWA's Demonstration Project No. 89 Quality Management can provide the vehicle to implement the activities under the initiative. This paper describes the development of these activities and their current status and plans.

There has been a conscious effort within the United States over the last decade to promote a correlation between American products and quality. In general, this effort has primarily been focused in the manufacturing industry. Congress has promoted the concept of American quality through the initiation of the Malcolm Baldrige National Quality Award (*1*). This award has received significant publicity from the recent recipients: Federal Express, Cadillac Motor Division, and IBM. Total quality management (TQM) is the subject of much discussion. Quality has become an important factor in maintaining global competitiveness.

Quality in the highway field is not new. Indeed, highway engineers have always been concerned with providing a quality roadway for the traveling public. The first use of formal quality assurance (QA) programs and specifications may be traced back to the AASHO Road Test in the early 1960s (*2*,p.3).

EARLY FHWA EFFORTS

In the 1970s, FHWA aggressively promoted QA programs through promotional and training efforts. Demonstration Project No. 2 made in-field side-by-side comparison of highway materials test results using innovative project sampling and testing programs and QA specifications. Early training included a course entitled Statistical Quality Control of Highway Construction for state highway agencies on the development of statistically based specifications. This effort was followed by Demonstration Project No. 42, a series of workshops that were designed to provide hands-on experience for state middle and upper management on development and implementation of a QA program.

The original FHWA policy guidance on QA programs stated: "The purpose of this directive is to establish a program to attain the widespread use of formal quality assurance techniques in highway construction by 1980" (*3*).

Early perceptions of QA significantly hampered progress in attaining widespread use. For example, the Associated General Contractors of America (AGC) in 1977 stated: "One of the major purposes of the use of statistics is to provide a formula for reduction of payment. . . ." (*4*,p.8) This perception was perhaps, and still may be, the greatest hindrance to expanding the use of quality management (QM) techniques today. In response to the AGC position, the then Chief of the Construction and Maintenance Division of FHWA, Sanford P. LaHue, stated (memorandum to Regional Administrators, Oct. 4, 1977): "These techniques are not limited to statistical specifications, but include such things as rapid testing procedures, improved process control, establishment of performance-related quality criteria, and the development of acceptance sampling and testing plans."

Today there is an even broader view of the subject of quality assurance. "Construction QM" is a broader term for the overall process of ensuring construction of quality products. It not only encompasses contractor process control and owner acceptance issues, including statistical quality control, but also such items as personnel qualifications, training, and certification programs; information management systems such as materials control systems and links to pavement management systems; performance-related specifications; innovative contracting practices to achieve quality; incentive-disincentive provisions to encourage quality attainment commensurate with the value received; performance recognition systems for quality projects and personnel; improved materials, tests, and equipment; and quality improvement techniques for both external and internal quality "customers."

CURRENT NEEDS

A few states have made significant progress in developing and implementing QM programs including specifications that rec-

Construction and Maintenance Division, Federal Highway Administration, 400 Seventh Street, S.W., Washington, D.C. 20590.

ognize materials and construction variability, revised sampling and testing programs, assignment of process control responsibility and activities to contractors, and training of contractor and state personnel.

Many of the state highway agency management personnel who directed the development and implementation of the early QM programs are now retiring and are being replaced by younger personnel who lack sufficient training in statistical quality control. In some states, personnel who were trained during the early FHWA efforts are now becoming managers and are expecting to change their QM procedures. At the same time, states are increasingly being pressured to adopt QM programs because of reductions in the level of state staffing, desires of the contracting industry, improved management practices, or all three.

In order to properly implement a successful QM program, there must be a substantial commitment to training long before such elements as new specifications or operating procedures are put into effect. Engineers, inspectors, construction supervisors, and construction workers all must be trained in their new responsibilities so that they can work as a team. Decisions need to be made on whether certification programs are going to be used and whether they will be in house, cooperative with industry, or by an outside agency such as the National Institute for Certification in Engineering Technologies (NICET).

Within the highway community, quality, and more particularly quality of the constructed product, has become a highly visible issue. Francis Francois, Executive Director of the American Association of State Highway and Transportation Officials (AASHTO), recently said in a letter to the Standing Committee on Highways (Nov. 8, 1990): "There is significant concern being expressed by some members of Congress about the quality of America's highway construction." In a comprehensive, four-part paper in this Record, Afferton et al. call for a national policy and leadership in QM.

OTHER MAJOR RELATED SUBJECT AREAS

Performance-Related Specifications

Although performance-related specifications (PRSs) do not have to be QM specifications, these are best suited to determine how much material is within specifications so that rational payment schedules can be developed. As mentioned before, in the past there was some sentiment that price adjustments were punitive in nature. To be equitable to all parties, price adjustments should be related to performance. Negative adjustments should rationally relate to the loss in service life and performance of the product. Relating specifications is a concept that was embraced by former Secretary of Transportation Samuel Skinner as part of his National Transportation Policy. One of the initiatives under that policy is that the U.S. Department of Transportation should "replace rigid standards and requirements with performance related criteria in Federal transportation programs" (*5*,p.44).

In 1990 a publication of the Transportation Research Board indicated that the highest priority research need in the nation was a project for development of PRSs. TRB said the objective would be "to improve quality control of highway construction by developing and implementing performance based specifications" (*6*,p.6). Other top priorities of the TRB study that relate to this initiative are

- Priority 3: Development of More Effective Rapid Test Methods and Procedures,
- Priority 5: Improving the Quality of Work on Highway Projects, and
- Priority 7: Responsibilities for Quality Management.

Currently, National Cooperative Highway Research Program (NCHRP) Project 20-5, Topic 23-05, on PRSs seeks to determine the extent to which construction and materials specifications have been rationally linked to performance data.

FHWA has research under way to address PRSs that may eventually lead to equitable and rational pay adjustment clauses. This includes research on pay adjustment provisions perhaps more likely to give quality results—price incentives for quality. FHWA has endorsed the use of incentives for improved quality provided they are based on readily measured characteristics that reflect improved performance. There is disagreement, however, among states about the factors to be used, the pay schedules, and how multiple factors are treated.

Innovative Contracting

In order to achieve quality results, there must be sufficient incentive or motivation by the provider to produce quality products. Critics argue that one of the obstacles in achieving quality in government contracting is the low-bid process. By law, construction contracts must be awarded to the lowest responsible bidder [Title 23, U.S. Code, Section 112, Par. b(1)]. Conversely, the law also mandates award of design contracts on the basis of qualifications and experience rather than cost [Title 23, U.S. Code, Section 112, Par. b(2)]. Although this may send a mixed signal that government is more concerned about cost than quality in construction contracts whereas quality is much more important than cost in design contracts, this is the current philosophy in the public-works sector of highway construction. Methods are being used or investigated to introduce quality into the construction bidding process. One method already mentioned is the use of incentives and disincentives for quality. If a good contractor is sure that a bonus can be earned, the bid price should reflect this fact. This then should give a "quality contractor" a slight price advantage in the low-bid process. Quality could also be entered directly into the bid. Factoring time estimated to complete a project as well as cost into the low bid, known as $A + B$ or multiparameter bidding, has been used to some extent in this country. It is conceivable that quality could somehow also be factored into the low-bid determination if an equitable method for quantifying quality, possibly through PRSs, could be developed.

At present the most common method of entering quality into the bidding process is the use of contractor prequalification procedures, in which the past work quality of contractors as well as their financial abilities are considered in determining which contractors are qualified to submit bids.

The use of warranties and guarantees has been the subject of much discussion lately. These have not been allowed in the past for federal-aid contracts (Title 23, Code of Federal Regulations, Part 635, Section 413), on the rationale that such

requirements would indirectly result in federal-aid participation in maintenance, which has long been prohibited. The use of warranties and guarantees is currently being studied by the General Accounting Office, however, as required under Section 1043 of the Intermodal Surface Transportation Efficiency Act of 1991. FHWA has been experimentally evaluating the use warranties and guarantees under FHWA Special Experimental Project 14, concentrating on warranty of products or features in such a manner to preclude any participation in routine maintenance.

AASHTO

The AASHTO Subcommittee on Construction has several activities under way concerning QM. The subcommittee has frequently discussed the incorporation of QM specifications into the AASHTO Guide Specifications for Highway Construction. Such incorporation was only one vote short of being accomplished in a balloting of the committee in 1978. Recently the western region of AASHTO, known as the Western Association of State Highway and Transportation Officials (WASHTO), created a task force and produced a set of guide QA specifications (*7*). The AASHTO Subcommittee on Construction plans to study the use of the WASHTO specifications as a basis for developing AASHTO guide specifications. The subcommittee is also developing an implementation guide for QM programs and specifications.

Another activity of the subcommittee is development of a QM data base to provide information on various aspects of QM, such as technician certification programs, use of contractor process control provisions, statistical quality control and acceptance, and use of incentives and disincentives for construction quality.

DEMONSTRATION PROJECT 89 WORKSHOP

FHWA initiated a demonstration project to gain top-level management support for QM principles, increase technical understanding, provide various references concerning specifications and implementation of QM programs, and, it is hoped, tie these separate efforts together.

As the first activity under Demonstration Project 89, FHWA sponsored a workshop consisting of top leaders in the QM field from state highway agencies, the construction industry, construction associations, academia, and FHWA. Approximately 30 leaders in the highway construction quality field were brought together on December 12–13, 1990, to discuss quality of the constructed product and to provide input into FHWA's role. The individuals attending this workshop were invited to represent a broad cross section of the highway industry.

On the first day of the workshop, a number of presentations were made on past and current activities regarding quality, followed by presentations on TQM and how much of this philosophy follows the concepts of QA in the highway industry. Other presentations and group discussions focused on specific elements of a QM program. In general, the presentations and group discussions emphasized the following key elements that must be included in a successful construction QM program:

- Absolute commitment of top management to provision of quality products;
- Programs for quality improvement and for quality assurance, both internal and external to the organization;
- Training and certification programs;
- Well-written, sound statistical specifications;
- Use of performance feedback information in evaluating and refining specifications;
- Involvement by industry in the development of quality management specifications;
- Use of rational and equitable incentive-disincentive provisions; and
- Provision of the necessary tools and resources.

Ken Afferton, Assistant Commissioner for Design and Right-of-Way of the New Jersey Department of Transportation, suggested that the following were needed on a national level to broaden proper use of QM techniques:

1. Education in QA and QM,
2. Improved national guidelines,
3. A national policy statement, and
4. An FHWA mandate similar to that for pavement management.

On the second day, small working groups discussed and later presented recommendations on future national activities concerning quality in the highway industry. The response of the workshop participants was that the needed emphasis of QM must go deeper than a demonstration project. There must also be a long-term commitment by FHWA through policy issuance, training, and technical support.

The workshop participants concluded that there is a need for the development of a "quality consciousness" within the highway community, but no agency or organization has been willing to initiate such action. The following is a composite of specific recommendations.

- A national initiative on quality is essential.
- Top management understanding of and commitment to quality products and delivery are critical.
- A national statement of policy should be a part of an initiative in order to show national commitment. It should be developed jointly by FHWA, AASHTO, industry, and academia.
- FHWA should affirm its commitment to quality and provide the needed leadership in developing a national initiative on quality.
- One part of the initiative should be a demonstration project focused on design, construction, materials, and maintenance quality. The focus should be broader than statistical quality control specifications.
- A major emphasis should be placed on partnership among designers, owners, contractors, and suppliers in achieving quality results.
- Technical skills and tools are essential and should be provided. These include certification programs, sound statistical

specifications, and a long-term commitment to technical training.

• A long-range plan for implementing a national initiative on quality should be developed and followed.

A smaller working group formulated the overall objectives of a quality initiative to emphasize that quality is integral to each element of the National Transportation Policy. Specific objectives would be to

1. Improve the technical quality and responsiveness to public needs of the nation's transportation systems;
2. Increase the strength and competitiveness of the U.S. transportation industry in the global marketplace through quality emphasis and improvement;
3. Advance the quality of transportation delivery systems through partnership efforts among FHWA, AASHTO, industry, and academia;
4. Maximize the use of the transportation investment through better system and product performance; and
5. Encourage technological developments and innovations through quality incentives.

An initial list of possible elements of a national initiative was also developed by the workshop participants. These were

- Marketing initiatives:
  - One- to two-day seminars for top management of FHWA and AASHTO and industry chief executive officers,
  - Mid-level manager awareness and implementation training,
- Improved NICET model and information series,
- QM and statistical training:
  - Contractor oriented,
  - State employee oriented,
- Repackaged and expanded availability of exciting reference materials and training,
- Documents containing recommendations on implementing QM,
- Development of easy, modern statistical tools.

The support for use of federal mandates was very mixed. Some argue that unless there is a federal mandate, progress will be slow and perhaps nonexistent. There is also the position that some degree of uniformity between states allows increased competition. In general, industry representatives were strongly in favor of mandates. States would be very concerned, however, about mandates' being too prescriptive.

The issue of federal mandates is always controversial. Even internal government opposition to the proposed mandatory governmentwide use of TQM brought about the demise of the proposed Office of Management and Budget Circular A-123 (*8*).

NATIONAL INITIATIVE

It was decided that a briefing paper would be prepared for upper FHWA management, to gain support for a national initiative and for development of a joint policy statement with AASHTO and industry regarding QM. FHWA developed an internal position paper that included the major recommendations of the workshop. This position paper was endorsed by Thomas Larson, Federal Highway Administrator, on March 26, 1991.

The workshop participants suggested establishing a panel of top management from FHWA, AASHTO, and various industry representatives to discuss the need, form, and content of a national policy on highway quality. A national initiative would be composed of many elements, some of which can be accomplished under the auspices of Demonstration Project 89 or under the training efforts of the National Highway Institute. A general framework of a national initiative will be established and Technical Advisory Committees (TACs) formed to correspond to the major elements of the initiative. These TACs will guide the development of each component of an overall program.

Francis Francois, Executive Director of AASHTO, presented this concept to the Standing Committee on Highways (SCOH) on June 9, 1991, and SCOH voted to

- Approve AASHTO's commitment to a Construction Quality Assurance Initiative with FHWA and the construction industry.
- Authorize the SCOH chairman to appoint a task force to serve as AASHTO's representatives on a joint AASHTO-FHWA-industry steering committee to guide the Construction Quality Assurance Initiative. The joint steering committee would
  - Develop a draft statement endorsing a construction quality initiative;
  - Guide development of seminars for top-level leaders from states and industry to further the understanding of the purposes, benefits, and techniques for improving construction quality; and
  - Provide suggestions and guidance to AASHTO's committees and member departments aimed at the overall improvement of highway construction quality (*9*).

Solicitation of industry support was through the joint committee meeting of AASHTO, the American Road and Transportation Builders Association, and the AGC in August 1991 and by letter from Francois. FHWA under Demonstration Project 89 is developing various contracts and work agreements for seminars, technical training, and preparation of various reference documents and tools.

A COMMON THREAD AND VEHICLE

The national initiative is needed to bring national attention to the quality of construction, further the use of construction QM, and help tie together the many efforts currently under way in QM. If it is to succeed, there must be a positive, unselfish partnership among FHWA, AASHTO, and the contracting industry. Political leaders, top managers, and industry leaders must be made aware of the potential benefits of a comprehensive QM program including sound design, construction, and maintenance practices; valid and effective statistical acceptance procedures; and wise infrastructure investment policies. Demonstration Project 89 can be the vehicle to bring the quality issue to national focus and deliver

many of the necessary products. This should also reaffirm FHWA's commitment toward promoting QM and quality of construction.

REFERENCES

1. Public Law 100–107. August 20, 1990.
2. *NCHRP Synthesis 65: Quality Assurance.* TRB, National Research Council, Washington, D.C., Oct. 1979.
3. *Federal Highway Program Manual 6–4–2–10* (rescinded). FHWA, U.S. Department of Transportation, May 27, 1975.
4. Statistically Oriented End-Result Specifications. Associated General Contractors of America, Aug. 1977.
5. *Moving America: New Directions, New Opportunities—A Statement of National Transportation Policy, Strategies for Action.* U.S. Department of Transportation, 1990.
6. Transportation Research Board. *Unpublished Report 45: Research and Development Program for Highway Construction Engineering Management.* Report FHWA-RD-90-034. FHWA, U.S. Department of Transportation, Feb. 1990.
7. *Model Quality Assurance Specifications.* WASHTO, Aug. 1991.
8. T. Shoop. Uphill Climb To Quality. *Government Executive*, March 1991, pp. 17–19.
9. *The AASHTO Journal*, Vol. 91, No. 23, June 14, 1991, p. 8.